

## FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA  
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCTCTTCCCG  
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG  
CTGCTGCCGCCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT  
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA  
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATC  
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA  
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT  
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC  
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA  
GGGCGACGGGTCTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG  
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC  
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT  
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG  
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG  
GGCTGCACAGGGGAAGGCCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA  
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG  
AAAAGTGTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTGACGGCTTCGAAGAAACG  
GAAGATGCCTGTGTGCCGCCGCGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT  
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC  
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC  
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA  
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG  
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC  
TTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA  
TCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC  
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTTCGGCGCAGCACCATGGCCTGAAAT  
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG  
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC  
ATCTCAATTAGTCAGCAACCCAGTTTT

U950401.1 "U950401.1"

## FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEKTLISKYESSEIRL  
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG  
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGECEVGWVLDE  
GACVDVDECAAEPPPCSAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT  
CVRKNENCYNTPGSYVCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

### **Signal peptide:**

amino acids 1-24

### **N-glycosylation sites.**

amino acids 190-194 and 251-255

### **Glycosaminoglycan attachment sites.**

amino acids 149-153 and 155-159

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 26-30

### **Casein kinase II phosphorylation sites.**

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343  
and 349-353

### **Tyrosine kinase phosphorylation site.**

amino acids 303-310

### **N-myristoylation sites.**

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and  
313-319

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 308-320

### **EGF-like domain cysteine pattern signature.**

amino acids 166-178

### **Leucine zipper pattern.**

amino acids 94-116

### FIGURE 3

U594001.1 "U594001.1"

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC  
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACA  
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG  
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCCGGAGGAGCGCCTTC  
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC  
CGGGCCGCCGAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA  
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT  
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC  
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCTTGGCGCTCCCTGGATA  
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA  
TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTGAAGT  
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT  
TCTTTAAACATGTCAACAAGCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT  
GAAAGACGCATCTGCGAGTGTCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG  
TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG  
GATTCTATGGAGTGAACGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC  
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG  
CAAATGCCCAACCCCTGTGCAAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT  
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT  
GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA  
TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC  
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG  
**TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCA  
TGTGTTGAATGTTCAAATAATGTTTATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT  
TCATTATAAATCACTGAGCTGATATTTACTCTTCTTTAAGTTTTCTAAGTACGTCTGTAG  
CATGATGGTATAGATTTTCTTGTTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA  
TCAGGTTAAAATTTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT  
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTTGGGCAAAAATGCGTAAGTCACAAGAAT  
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT  
GTTTGTACATTTTTTAAAAATTGCTCTTAATTTTTTAACTCTCAATACAATATATTTTGACC  
TTACCATTATTCCAGAGATTGAGTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATT  
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT  
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT  
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGAAGCTTGGC  
CGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

## FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESELYLWIDAHQARVLIGFEEDILIVSEGK  
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL  
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC  
RNGGFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST  
TCFNNGGTCTFYPGKICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC  
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP  
PESNYIW

### **Signal peptide:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 88-92, 245-249

### **Casein kinase II phosphorylation site.**

amino acids 319-323

### **Tyrosine kinase phosphorylation site.**

amino acids 370-378

### **N-myristoylation sites.**

amino acids 184-190, 185-191, 189-195, 315-321

### **ATP/GTP-binding site motif A (P-loop).**

amino acids 285-293

### **EGF-like domain cysteine pattern signature.**

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

[illegible][illegible]

## **FIGURE 6**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC  
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNR  
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

### **N-glycosylation site.**

amino acids 78-82, 161-165

### **Casein kinase II phosphorylation site.**

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,  
411-415

### **N-myristoylation site.**

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,  
269-275, 378-384, 442-448

### **Amidation site.**

amino acids 26-30, 318-322

### **Eukaryotic thiol (cysteine) proteases histidine active site.**

amino acids 398-409

## **FIGURE 7**

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTTCGTTGTCTCAAATCCA  
ATTCTCTTGGGACACATNACGCCTGTCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC  
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCTTGCCTCGCCGAGG  
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC  
CCCCCTGTATGATGCACAGCCGAGCCATGGGTCTGGGGCAAGCGCCAGGCCACTGCCCCTGCGC  
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC  
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG  
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG  
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

0590401.07101

## **FIGURE 8**

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT  
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG  
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG  
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT  
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT  
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG  
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG  
GCCCAGTGGGAGCCTGTCTTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT  
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA  
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC  
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT  
CCCTGCCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA  
GGGGACAGGCACTCAGGAGGGGCCAGTTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA  
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA  
GGGGCCAGGCCTCACATTCTGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT  
AATAAACACCTGTTGGATAAGCCAAAAAA

05504011-07101



DATE	TIME	LOCATION	REMARKS
1944	10:30	1000	1000
1944	11:00	1000	1000
1944	11:30	1000	1000
1944	12:00	1000	1000
1944	12:30	1000	1000
1944	13:00	1000	1000
1944	13:30	1000	1000
1944	14:00	1000	1000
1944	14:30	1000	1000
1944	15:00	1000	1000
1944	15:30	1000	1000
1944	16:00	1000	1000
1944	16:30	1000	1000
1944	17:00	1000	1000
1944	17:30	1000	1000
1944	18:00	1000	1000
1944	18:30	1000	1000
1944	19:00	1000	1000
1944	19:30	1000	1000
1944	20:00	1000	1000
1944	20:30	1000	1000
1944	21:00	1000	1000
1944	21:30	1000	1000
1944	22:00	1000	1000
1944	22:30	1000	1000
1944	23:00	1000	1000
1944	23:30	1000	1000
1944	00:00	1000	1000
1944	00:30	1000	1000
1944	01:00	1000	1000
1944	01:30	1000	1000
1944	02:00	1000	1000
1944	02:30	1000	1000
1944	03:00	1000	1000
1944	03:30	1000	1000
1944	04:00	1000	1000
1944	04:30	1000	1000
1944	05:00	1000	1000
1944	05:30	1000	1000
1944	06:00	1000	1000
1944	06:30	1000	1000
1944	07:00	1000	1000
1944	07:30	1000	1000
1944	08:00	1000	1000
1944	08:30	1000	1000
1944	09:00	1000	1000
1944	09:30	1000	1000
1944	10:00	1000	1000
1944	10:30	1000	1000
1944	11:00	1000	1000
1944	11:30	1000	1000
1944	12:00	1000	1000
1944	12:30	1000	1000
1944	13:00	1000	1000
1944	13:30	1000	1000
1944	14:00	1000	1000
1944	14:30	1000	1000
1944	15:00	1000	1000
1944	15:30	1000	1000
1944	16:00	1000	1000
1944	16:30	1000	1000
1944	17:00	1000	1000
1944	17:30	1000	1000
1944	18:00	1000	1000
1944	18:30	1000	1000
1944	19:00	1000	1000
1944	19:30	1000	1000
1944	20:00	1000	1000
1944	20:30	1000	1000
1944	21:00	1000	1000
1944	21:30	1000	1000
1944	22:00	1000	1000
1944	22:30	1000	1000
1944	23:00	1000	1000
1944	23:30	1000	1000
1944	00:00	1000	1000
1944	00:30	1000	1000
1944	01:00	1000	1000
1944	01:30	1000	1000
1944	02:00	1000	1000
1944	02:30	1000	1000
1944	03:00	1000	1000
1944	03:30	1000	1000
1944	04:00	1000	1000
1944			

**Signal peptide:**

N-glycosylation site.

**Casein kinase II phosphorylation site.**

N-myristoylation site.

Amidation site.

amino acids 70-74

Figure 1 displays 12 histograms, labeled  $x_0$  through  $x_{11}$ , showing the distribution of the number of non-zero elements in the vector  $x_k$ . The x-axis represents the number of non-zero elements (0 to 10), and the y-axis represents the count (0 to 10). The distributions are roughly bell-shaped and centered around 5, with the peak count increasing from 10 at  $x_0$  to 12 at  $x_{11}$ .

[illegible]

## **FIGURE 11**

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMATRQGRPRQASRSRQONQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

**Signal peptide:**

amino acids 1-22

**N-glycosylation site.**

amino acids 9-13, 126-130

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 60-64

**Casein kinase II phosphorylation site.**

amino acids 65-69

**Tyrosine kinase phosphorylation site.**

amino acids 39-48, 89-97

**N-myristoylation site.**

amino acids 69-75, 188-194

**Amidation site.**

amino acids 58-62

**HBGF/FGF family signature.**

amino acids 103-128

## FIGURE 12

U544411.1 "U544411.1

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG  
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC  
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT  
CATTTTGATTTTGCTGTTTATTTTTTTTTTTCTTTTCTTTTCCCACCACATTGTATTTTAT  
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT  
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC  
CTAGTGTGTGCCGCTGCGACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG  
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC  
TGGATTTCTCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC  
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC  
AATATTCAGACCATTTTACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT  
GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC  
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC  
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCTATATCCGACATGGCCTTCAGAA  
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG  
AGGGCACCTTCAGCCATCTCACCAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC  
CACCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT  
AAACCACATTCTTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA  
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC  
ACTGCTCGGAATAACCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA  
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG  
GGATGGCCGTGAGGGAATTAATATGAATCTTTTGTCTGTCCCACCACGACCCCGGCCTG  
CCTCTCTTACCCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT  
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAAACTTCCCACGATTC  
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC  
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCCTGATGGCATA  
CAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTTACGAGCGCATAG  
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT  
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC  
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT  
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAGGGGCGCTACACCTC  
CCAGAAGTGGAAATACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA  
AGGACAACCTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT  
CAACTCCTTAAAGGAGATTTTCAAGTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA  
CACAGACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC  
ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA  
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT  
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG  
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTTAAATCTT

## **FIGURE 13**

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP  
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI  
SRAALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR  
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD  
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN  
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMVRELNMNLLSCPTTTPGLPLFTP  
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND  
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL  
DAFNRYAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL  
LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG  
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

**Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 542-561

**N-glycosylation site.**

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

**Casein kinase II phosphorylation site.**

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

**Tyrosine kinase phosphorylation site.**

amino acids 319-328

**N-myristoylation site.**

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,  
522-528, 545-551, 633-639

**Amidation site.**

amino acids 581-585

**Leucine zipper pattern.**

amino acids 164-186

**Phospholipase A2 aspartic acid active site.**

amino acids 39-50

650  
640  
630  
620  
610  
600  
590  
580  
570  
560  
550  
540  
530  
520  
510  
500  
490  
480  
470  
460  
450  
440  
430  
420  
410  
400  
390  
380  
370  
360  
350  
340  
330  
320  
310  
300  
290  
280  
270  
260  
250  
240  
230  
220  
210  
200  
190  
180  
170  
160  
150  
140  
130  
120  
110  
100  
90  
80  
70  
60  
50  
40  
30  
20  
10  
0

# FIGURE 14

U55443.1 "U55443.1"

ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAAGCTGGGGCTCCGTCTCGCCTCCACGAGCG  
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCCTGCC  
TCGCTTCCCAGGCGCCGCGGCTGCAGCCTTGCCCCCTCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT  
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTACGTGGGAGGTCCATCT  
CTAGGGGCAGACACGCTCGGACCCACCCGAGACGGCCCTTCTGGAGAGTTCTGTGAGAACAGCGGGCAGACC  
TGGTTTTTCATCATTTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTTCATCGTGGACA  
TCTTGCAATTCTTGGACATTGGTCCTGATGTCAACCGAGTGGGCTGCTCCAATATGGCAGCACTGTCAAGAATG  
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCATCTGTCCACGG  
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCGCGCCCTGA  
GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA  
AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG  
GGAGTGAGCCCCATGAGGACCATGTCTTCTTGTGGCCAATTTAGCCAGATTGAGACGCTGACCTCCGTGTTCC  
AGAAGAAGTTGTGCACGGCCACATGTGCAGCACCTTGGAGCATAACTGTGCCACTTCTGCATCAACATCCCTG  
GCTCATACGTCTGCAGGTGCAAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT  
GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG  
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCCTCAGAAAACACGGATGTGAAC  
ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAA  
CGTGACAAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCCTCAACATGGAGGAGAGCT  
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC  
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCTTCGTCTGCCAGTGCTCAGAAGGCTTCC  
TCATCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTACTGCCTGTGAGTGACCATGGTTGTGAATACTCCT  
GTGTCAACATGGACAGATCCTTTGCCGTGTCAGTGTCTGAGGGACACGTGCTCCGACGATGGGAAGACGTGTG  
CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGTAAGCAGTGAAGATTCTGTTGTG  
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG  
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACAGTGCGAGTGCTTGGAGGGATTCCGGCTCG  
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAATCAACCCACCATGGCTGCCAACACATTTGTGTTA  
ATAATGGGAATTCCTACATCTGCAATGCTCAGAGGGATTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT  
GCACCTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAAGAGTCTTGAGAAGAGAATTTTGAGGTCTG  
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT  
ATTCCACACAGGTCCACACAGAGTTCACCTCTGAGAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC  
ACATGAAATACATGGGAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCCAAG  
GAGAAGGGGCCAGGCCCCCTTCCACAAGGGTGCCGAGAGCAGCCATTGTGTTTACCGACGGACGGGCTCAGGATG  
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG  
AGGAGGAACATAAAGAGATTGCCTCTGAGCCCAACAAACAAGCATCTCTTCTATGCCGAAGACTTCAGCACAAATGG  
ATGAGATAAGTGAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG  
CAGGGGAACCTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT  
CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT  
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACAGATCAATGCAAATGTGAAAACCTTATAATGT  
TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC  
TGGAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT  
GAACGCAGTGACAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAAACAATCAGTACTGA  
GAAACCTGGTTTGGCACAGAACAAGACAAGAAGTATACACTAAGTTGTATAAATTTATCTAGGAAAAAATCCT  
TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAAATATACTGTGGACAC  
AACTTGCTTCTGCCTCATCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT  
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG  
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAATAATTCACCACTTCAG

## FIGURE 15

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII  
DSSRSVINTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV  
KMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD  
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN  
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA  
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINICALNKP GC  
EHECVNMEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI  
NEDLKTC SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTC AKLDSCALGDHGCE  
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA  
EDGKR<sup>1</sup>RRKDVC<sup>2</sup>STHHGCEHICVNNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID  
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA  
HMKYMGKGSMTGLALKHMFERSFTQEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN  
GITMYAVGVGKAIEEELQEIASEPTNKHLYAEDFSTMDIEISEKLKKGICEALEDSDGRQDS  
PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL  
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

**Signal peptide:**

amino acids 1-23

**N-glycosylation site.**

amino acids 221-225

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 115-119, 606-610, 892-896

**Casein kinase II phosphorylation site.**

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,  
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,  
781-785, 819-823, 866-870

**N-myristoylation site.**

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,  
639-645, 690-696, 752-758, 792-798

**Amidation site.**

amino acids 314-318, 560-564, 601-605

**Aspartic acid and asparagine hydroxylation site.**

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,  
540-546, 581-587

CC-BY-NC-ND 4.0 International license

## FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCG  
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCCTGACTCCGTCCCGGCCAGGGAGGGC  
CATGATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA  
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG  
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC  
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC  
AGGTGTTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC  
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA  
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT  
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT  
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA  
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG  
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC  
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC  
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC  
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC  
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT  
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT  
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT  
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTTCTTCTCTGGCTTGAGCCG  
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC  
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCTATAAGGGTCACCTCTAGCAC  
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC  
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA  
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG  
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGC  
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT  
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC  
TTATTAAAACTAACATGAAATATGTGTTGTTTTTCATTTGCAAATTTAAATAAAGATACATAA  
TGTTTGTATGAAAAA

U590461.07101



[illegible]

**Signal peptide:**

**Transmembrane domain:**

**N-glycosylation site.**

N-myristoylation site.

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 164-175

# FIGURE 18

U55640.1 "U55640.1

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTTCCACTTTG  
 TTGAATTGTTCTTACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA  
 AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTACAA  
 TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC  
 ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA  
 AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT  
 TAGATAATGTCTGTATAGCTGCAAATATTAATAAACTTTAAACAAAATCAGATCCATAAAA  
 GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA  
 TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA  
 CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT  
 AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC  
 ACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC  
 AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGTAT  
 TCATATAACATGAAACATATTCTCCTCATATGAATATGGATGGAGACTACATAAATATATT  
 TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTTATATTATA  
 AGAGTATTGGTCCTTTGCTTTTCATCATCTGACAACCTTCTTATTGAAACCTCAAAATTATGAT  
 AATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTTCAGTCTCAATGAGCTCAAACCC  
 ACCCACATTATATGAACCTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA  
 GGTATAGGAGTCTATGTGCATTTTGGAACTTACTCACCTGATACCATGAATGGCAGCTGGTCT  
 TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT  
 GACACATTTTGAATTTTGTATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC  
 TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT  
 ACCTTCTGGTTCTTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG  
 TAGCCTATTTCTTGCTGAACCTTGTTTTCTTGTGGGATCAATACAAATACTAATAAGCTCT  
 TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGC  
 ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA  
 CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC  
 TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCCGAAAACAACCTTTATT  
 TGGAGTTTTATAGGACCAGCATGCCTAATCATTTCTTGTTAATCTCTTGGCTTTTGGAGTCAT  
 CATATACAAAGTTTTTTCGTACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA  
 TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCCCTTCTCGGCACCACCTGGATCTTT  
 GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC  
 TTTCCAGGGGATGTTCAATTTTTTTATTCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT  
 ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGGTAAACATAGAGAATG  
 GTGGATAATTACAACCTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA  
 TGACTCATCAAATTATCCAATTATTAACCTAGACAAAAAGTATTTTAAATCAGTTTTTCT  
 GTTTATGCTATAGGAACCTGTAGATAATAAGGTAAATTTATGTATCATATAGATATACTATGT  
 TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT  
 CTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTCTAACACGAGAAGTATATGAA  
 TGTCTGAAGGAAACCACTGGCTTGATAATTTCTGTGACTCGTGTGCTTTGAAACTAGTCC  
 CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA  
 ATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGAC  
 TAGCTGAGAAATTGTTGACATAAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA  
 TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT  
 TTTCTAATATTCTAAAAAAGGTTTACCTCCACAAATTGAAAAAAGGTTTACCTCCACAAATGAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 19**

MKRLPLL VVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC  
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA  
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL  
SNSTLTETVKT VNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT  
NSTDIALKVFFFDSYNMKHIHPHMNMGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS  
SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTD RYRSLCAF  
WNYSPTMTNGSWSSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG  
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFC SIIAGL  
LHYFFLA AFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFS AALGYRYYGT  
TKVCWLSTENNF IWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA  
LALLFLLGTTWIFGVLHVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV  
PCCFGCLR

### **Signal peptide:**

amino acids 1-19

### **Transmembrane domain:**

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,  
648-664

### **N-glycosylation site.**

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,  
188-192, 249-253, 381-385, 395-399

### **Glycosaminoglycan attachment site.**

amino acids 49-53

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 360-364

### **Casein kinase II phosphorylation site.**

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,  
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,  
346-350, 608-612

### **Tyrosine kinase phosphorylation site.**

amino acids 36-44, 669-677, 670-678

### **N-myristoylation site.**

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,  
434-440, 480-486, 521-527

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 75-87

[illegible]

**THE UNIVERSITY OF CHICAGO**

[illegible]

[illegible]

**Signal peptide:**

**Casein kinase II phosphorylation site.**

N-myristoylation site.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59



[illegible]

amino acids 1-20

amino acids 130-144, 238-258

amino acids 98-102, 187-191, 236-240, 277-281

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

## amino acids 226-230



	9	8	7	6	5	4	3	2	1	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	-25	-26	-27	-28	-29	-30	-31	-32	-33	-34	-35	-36	-37	-38	-39	-40	-41	-42	-43	-44	-45	-46	-47	-48	-49	-50	-51	-52	-53	-54	-55	-56	-57	-58	-59	-60	-61	-62	-63	-64	-65	-66	-67	-68	-69	-70	-71	-72	-73	-74	-75	-76	-77	-78	-79	-80	-81	-82	-83	-84	-85	-86	-87	-88	-89	-90	-91	-92	-93	-94	-95	-96	-97	-98	-99	-100	-101	-102	-103	-104	-105	-106	-107	-108	-109	-110	-111	-112	-113	-114	-115	-116	-117	-118	-119	-120	-121	-122	-123	-124	-125	-126	-127	-128	-129	-130	-131	-132	-133	-134	-135	-136	-137	-138	-139	-140	-141	-142	-143	-144	-145	-146	-147	-148	-149	-150	-151	-152	-153	-154	-155	-156	-157	-158	-159	-160	-161	-162	-163	-164	-165	-166	-167	-168	-169	-170	-171	-172	-173	-174	-175	-176	-177	-178	-179	-180	-181	-182	-183	-184	-185	-186	-187	-188	-189	-190	-191	-192	-193	-194	-195	-196	-197	-198	-199	-200	-201	-202	-203	-204	-205	-206	-207	-208	-209	-210	-211	-212	-213	-214	-215	-216	-217	-218	-219	-220	-221	-222	-223	-224	-225	-226	-227	-228	-229	-230	-231	-232	-233	-234	-235	-236	-237	-238	-239	-240	-241	-242	-243	-244	-245	-246	-247	-248	-249	-250	-251	-252	-253	-254	-255	-256	-257	-258	-259	-260	-261	-262	-263	-264	-265	-266	-267	-268	-269	-270	-271	-272	-273	-274	-275	-276	-277	-278	-279	-280	-281	-282	-283	-284	-285	-286	-287	-288	-289	-290	-291	-292	-293	-294	-295	-296	-297	-298	-299	-300	-301	-302	-303	-304	-305	-306	-307	-308	-309	-310	-311	-312	-313	-314	-315	-316	-317	-318	-319	-320	-321	-322	-323	-324	-325	-326	-327	-328	-329	-330	-331	-332	-333	-334	-335	-336	-337	-338	-339	-340	-341	-342	-343	-344	-345	-346	-347	-348	-349	-350	-351	-352	-353	-354	-355	-356	-357	-358	-359	-360	-361	-362	-363	-364	-365	-366	-367	-368	-369	-370	-371	-372	-373	-374	-375	-376	-377	-378	-379	-380	-381	-382	-383	-384	-385	-386	-387	-388	-389	-390	-391	-392	-393	-394	-395	-396	-397	-398	-399	-400	-401	-402	-403	-404	-405	-406	-407	-408	-409	-410	-411	-412	-413	-414	-415	-416	-417	-418	-419	-420	-421	-422	-423	-424	-425	-426	-427	-428	-429	-430	-431	-432	-433	-434	-435	-436	-437	-438	-439	-440	-441	-442	-443	-444	-445	-446	-447	-448	-449	-450	-451	-452	-453	-454	-455	-456	-457	-458	-459	-460	-461	-462	-463	-464	-465	-466	-467	-468	-469	-470	-471	-472	-473	-474	-475	-476	-477	-478	-479	-480	-481	-482	-483	-484	-485	-486	-487	-488	-489	-490	-491	-492	-493	-494	-495	-496	-497	-498	-499	-500	-501	-502	-503	-504	-505	-506	-507	-508	-509	-510	-511	-512	-513	-514	-515	-516	-51
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

AAAAACCACGAAGGAACTACTCCAAAAATGAAC

## **FIGURE 26**

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT  
FPARLPANTQILLQLQTNNAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE  
ENKLTPEKCLSELNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA  
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL  
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPD  
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSPNIRC  
DCVIRWMNMNKTNIIRFMEPDSLFCVDPPEFQGNVRQVHFRDMMEICLPLIAPESFPSNLNV  
EAGSYVSFHCRTAEAPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTC  
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFAV  
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE  
YEKNNTTTLMACLGGLLGIIGVICLISCLSPMNCDGGHSYVRNYLQKPTFALGELYPPLIN  
LWEAGKEKSTSLKVKATVIGLPTNMS

### **Signal sequence:**

amino acids 1-22

### **Transmembrane domain:**

amino acids 633-650

### **N-glycosylation site.**

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,  
608-612, 624-628, 625-629

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

### **Tyrosine kinase phosphorylation site.**

amino acids 570-579

### **N-myristoylation site.**

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,  
491-497, 492-498, 634-640, 702-708

### **Cell attachment sequence.**

amino acids 277-280

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG  
 CTGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC  
 CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT  
 GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTGTCTTA  
 TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG  
 GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA  
 AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG  
 ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT  
 GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG  
 TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC  
 ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCAC  
 AACGTGATCTGTAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC  
 CAACGACGCTGACCTTTGTAACCTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA  
 TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAATCAGGAG  
 GATGCCCCGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA  
 ACCTGATGATATTAGCACTGTGGTATTAGTGTCCAACTGACTGTCATTGAGAAAGAAAGAAA  
 GTAGTTTTCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA  
 ACTTTGTATTTTCAGTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA  
 TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT  
 AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT  
 AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAAAACA

## **FIGURE 28**

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL  
PPETVLLYLDSDNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNR  
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVIKTSVLDEHAGRPFL  
NAANDADLCNLPKKTDDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK  
ADEPDDISTVV

### **Signal sequence:**

amino acids 1-33

### **Transmembrane domain:**

amino acids 205-220

### **N-glycosylation site.**

amino acids 47-51, 94-98

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 199-203

### **Casein kinase II phosphorylation site.**

amino acids 162-166, 175-179

### **N-myristoylation site.**

amino acids 37-43, 45-51, 110-116

T  
E  
X  
T  
F  
I  
L  
E

[illegible]

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCGAGGTTGAGCAAGAGGATGCTGGCGGGG  
 GGCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT  
 GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCCGCCCCGCTGCGAGTGCTCCGCCCAGG  
 ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG  
 ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG  
 CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG  
 CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC  
 CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT  
 CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG  
 ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG  
 ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT  
 CATCGTCTTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC  
 TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC  
 TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC  
 CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCTCCTCAACCTCTCCTACAACCCCATCA  
 GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC  
 GGGCAGCTGGCCGTGGTGGAGCCCTATGCCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA  
 TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG  
 AGAACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTGCGCTCCTGTGGGTGTTCCGG  
 CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCACGTGCGCCACGCCCGAGTTTGTCCA  
 GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG  
 CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGGCCACACGGTGCGAGTTT  
 GTGTGCCGGGCGGATGGCGACCCGCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT  
 GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT  
 ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC  
 TCCATGCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA  
 GACCTTCGCTTTCATCTCCAACAGCCGGGCGAGGGAGAGGGCCAACAGCACCCGCGCCACTG  
 TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC  
 CTGGGCGTCGTCTCTCTGCTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC  
 AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG  
 ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGAGGCCGGGGCGGGGGGACCCCCG  
 GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCTTCCCACCTC  
 CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCCTCCGTCCCCTGCTGCCCCCG  
 CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA  
 CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA  
 ATAATTCAATAAAAAAGTTACGAACTTTCTCTGTAACCTTGGGTTTCAATAATTATGGATTTT  
 TATGAAAACCTTGAAATAATAAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 30**

MQVSKRMLAGGVRSMPSPLLACWQPIILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCF  
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFNLRTL  
GLRSNRLKLIPLGVFTGLSNLTQDISENKIVILLDYMFDLYNLKSLEVGDNLDVYISHRA  
FSGLNSLEQLTLEKCNLTSLIPTALSHLHGLIVLRLRHLNINAI RDYSFKRLYRLKVLEISH  
WPYLDTMTPNCLYGLNLTSLITHCNLTAVPYLAVRHLVYL RFLNLSYNPISTIEGSM LHEL  
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLT TLEESVFH SVGNLETLILDSNPLA  
CDCRLLWVFRRRWRLNFNRRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV  
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGR LTVFPDGTLEVRYAQVQDNGTYL  
CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI  
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 556-578

### **N-glycosylation site.**

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,  
492-496, 505-509, 526-530, 542-546

### **Casein kinase II phosphorylation site.**

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,  
607-611

### **Tyrosine kinase phosphorylation site.**

amino acids 590-598

### **N-myristoylation site.**

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,  
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 32-43

Downloaded from www.sciencedirect.com

[illegible][illegible]

[illegible]

**Signal sequence:**

Transmembrane domain:

N-glycosylation site.

**Casein kinase II phosphorylation site.**

N-myristoylation site.

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18



## FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCTCTGG  
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTTCGCTG  
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT  
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA  
TCTGTTCTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA  
AGTCTGCAGCGTTTCACTGCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC  
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG  
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAGG  
CTGCACATCAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA  
TCTGGAATATCTCCAGGCTGATTTTAAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG  
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTACCTGCCAAC  
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC  
CTATGAGGAGGTCTTGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCTT  
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC  
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC  
CACCGAACAGGACTTGTGTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG  
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG  
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT  
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCTTAGCTAACA  
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC  
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA  
GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA  
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG  
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA  
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA  
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTCTCAACAACAACCTG  
CTGAGGTCCCTGCCTGTGGACGTGTTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA  
CAATTACTTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA  
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA  
GAACGCTTGGGTTCGGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT  
TAGAAAGGATTTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT  
CGCCACGTAACTTCGCACAGTAAAAACAGCACTGGGTGGCGGAGACCGGGACGCACTCC  
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT  
GTTTGTACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTATTATCCTGAGGAACCGAAAGC  
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT  
GACTCTTCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCACAGAGTGTATGACTG  
TGGCTCTCACTCGCTCTCAGACTAAGACCCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG  
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG  
CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT  
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA  
GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG  
AACTGACAGTGCCCTCGCCCTCGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC  
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCTGTGGATTAG  
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT  
TGTAATAAGTAACTTTGACTTCTGAC

[illegible]

**Signal sequence:**

amino acids 337-348

[illegible]

AGTCGACTGCGTCCCCGTGTACCCGGCGCCAGCTGTGTTCTTGACCCCAAGAATAAAGTACAGGGC  
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAAACTGTTGGC  
CGCTGGGCCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG  
AAGGGAGGGGGAAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC  
AGCTCTGCGTCTCTGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCCTCAG  
AGAATGAGGCCGCGGCTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCCGCGG  
CGGCGAAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCCTCGGGGGCCTGCTACAGCCTGC  
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC  
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG  
GCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT  
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCTCCTCCGACCCCGCGGTCTC  
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC  
GGTACTCCAGGCCACCGGTGGGGTTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC  
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCTGCGCCGCGCCCCGGG  
GCCGCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG  
TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA  
TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC  
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG  
CTTTGCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA  
GTGGGGAAGGACAGCCGACCCCTTGGGGGGGACCGGGGTGCCACCAGGCGCCCGCGGCCACT  
GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC  
ACCACTTGTCCCTGAACAAGACAATTTCAGTAACATCTATTCTGAGATTCTTCGATGGGGAT  
CACAGAGCACGATGTCTACCCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC  
CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT  
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT  
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCCTCTTCCAGCCA  
AGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC  
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG  
CAGAGGGTGCTTGCTGGCGGAGTCCCTCCTTGCTCTAGTGATGCATAGGGAAACAGGGGA  
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGGAACCAAGAGGAACTTAC  
TTGTGTAAGTACAAATTTCTGCAGAAATCCCCCTTCTCTAAATTTCCCTTTACTCCACTGAG  
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA  
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTTCGGAGAA  
TTTGAGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA  
TAAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTTCAGGCTAGGAGTAT  
ATTGGTTTCGAAATCCCAGGGAAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

## **FIGURE 36**

MRPAFALCLLWQALWPGPGGGGEHPTADRAGCSASGACYSLHHATMKRQAEEACILRGGALS  
TVRAGAE LRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE  
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA  
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG  
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA  
TSPVPQRTWPPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP  
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR  
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 399-418

### **N-glycosylation site.**

amino acids 189-193, 381-385

### **Glycosaminoglycan attachment site.**

amino acids 289-293

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 98-102, 434-438

### **Casein kinase II phosphorylation site.**

amino acids 275-279, 288-292, 342-346, 445-449

### **N-myristoylation site.**

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,  
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,  
477-483

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 262-274

Downloaded from www.jstor.org

## FIGURE 37

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG  
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC  
CCCCGGCTACCTGGGCGGCCCGCGGGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGCAGCCGA  
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA  
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCCGAGCGTCTACGCTGCCATGA  
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG  
CAGCAGTCCCCAGAGAGACCTGTTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT  
TATTGGCAGTGAAGGTTTTCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA  
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC  
CTGTGCCGCTATGACTTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTTGGCCG  
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA  
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA  
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC  
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC  
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC  
CGATATGATTATGTGGCTGTGTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA  
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT  
TTTTATCAGACTTAAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA  
AACTGCCTACAACACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA  
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT  
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAAACCATCACTCGCGATGGGAGTTTG  
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTTCAGCAGGCGGG  
CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC  
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGC AAAATCATGCCAAACAGC  
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG  
TTAAACAGTGAACGTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC  
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAAGATGG  
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAAGTTCTT  
TGCTGCTGTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA  
AGCTAAAAGTGTCAAGCGTTGACAGCTTGGAAGCGTTTATTTATACATCTCTGTAAAAGGAT  
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTAGAAGTGCAATATTTATAGT  
GTTATTTGTTTACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTCTA  
AATCAATGCTTAATAAAAATATTTTAAAGGAAAAAAAAAAAAA

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG

[illegible]

**Signal sequence:**

N-glycosylation site.

Casein kinase II phosphorylation site.

**Tyrosine kinase phosphorylation site.**

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,  
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

## FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCCACGGCGCCCGCGGGCTGGGGCGGTGCTTCTT  
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCGAAGG  
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCTCAACCTCCCAGGACCTATCTGG  
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG  
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG  
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCTG  
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT  
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCAACAAGCAGCAGGAGGCCCCGGACCTCTTCC  
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC  
CTTCCCTGTCCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG  
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCCTGTGGCC  
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT  
TTTGGCCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGG  
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT  
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG  
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA  
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA  
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG  
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC  
AGAAGACGAGTTGGTGGTGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA  
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG  
ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA  
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC  
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTTGAGAGTGGGGTAAGCACCCCTACCTG  
CCTTACAGAGCAGCCCAGGTACCCAGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC  
CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA  
ATTTCAAAAGTTTTTCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG  
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT  
TCTGTGTTACCCACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA  
AAGGTCTTGGAAGTTAAAAA

## **FIGURE 40**

MAPWPPKGLVPAVLWGLSLFLNLPGPiWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER  
TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLLELSEELVESWWFHKQ  
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDQCAG  
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPESNCLQCKKGWALHHLKCVDIDE  
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE  
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDLVVLQQMFFG  
IIICALATLAAKGDVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

### **Signal sequence:**

amino acids 1-29

### **Transmembrane domain:**

amino acids 372-395

### **N-glycosylation site.**

amino acids 79-83, 205-209

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 290-294

### **Casein kinase II phosphorylation site.**

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

### **N-myristoylation site.**

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,  
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,  
326-332, 372-378, 395-401

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 321-333

### **EGF-like domain cysteine pattern signature.**

amino acids 181-193

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA  
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCCTGGCCAGCCCCGGG  
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT  
GCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGGCCAGT  
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTTCAGCCAGAGC  
TTCCGAGAGGTGGCCGGCAGGTTCTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG  
CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG  
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC  
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA  
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA  
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG  
GAGCATCTGGGCCCCGCTGGCGTCCGGCGCCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC  
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG  
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG  
ATGTACATTGACCTGCAGGGGATGAAGTGGGCGGAGAACTGGGTGCTGGAGCCCCCGGGCTT  
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC  
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC  
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA  
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATTAGGCGCCTAGTG  
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTTCGAGGGTACCAGGAGAGCTG  
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT  
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA  
GAGCCCTTGCTCAGTTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT  
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTGCTATTGTTTACTTGTCTGTGAC  
TGGATCTGGGCTAAAGTCCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT  
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT  
AAAA

CGGAGGCTT

## **FIGURE 42**

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV  
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP  
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF  
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ  
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF  
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 158-162

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 76-80

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

### **N-myristoylation site.**

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

### **Amidation site.**

amino acids 74-78

### **TGF-beta family signature.**

amino acids 282-298

U56001-0101

[illegible]

GTCTGTTCCAGGAGTCTTTCGGCGGCTGTTGTGTCAAGTGGCCTGATCGCGAATGGGGACAAA  
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG  
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT  
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA  
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG  
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC  
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT  
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG  
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT  
GGGATAGTGATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCTT  
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA  
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT  
GTGGAGCGGAATGTGGGGGTCTATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT  
CTTGTTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA  
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG  
ACCTCGTCATTCTGTTGTGAGCCTGGTTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT  
CAGGTGCTACCGGACTCTGGCCCCCTGATGTCTGTAGTTTACAGGATGCCTTATTTGTCTTC  
TACACCCACACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC  
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTGTTTAAA  
GTGTTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC  
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGCAGGAATCTGCACTCAACTGCCCACCTGGC  
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTTACTGAC  
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG  
TGATGACACTGGGGTCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG  
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT  
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG  
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTG  
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA  
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC  
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLS CAYSGFSSPRVEW  
KFDQGD TTRLVCYNNKITASYEDRV TFLPTGITFKSVTREDTGTYT CMVSEEGGNSYGEVKV  
KLIVLVPPSKPTVNI PSSATIGNRAVLTCSEQDGSPPSEYTWFKDGI VMPTNPKSTRAFSNS  
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGT PMTSNAVRMEAVERNVGVIVA AVLVTLI  
LLGILVFGIWFAYS RGHFDR TKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

### **Signal sequence:**

amino acids 1-27

### **Transmembrane domain:**

amino acids 238-255

### **N-glycosylation site.**

amino acids 185-189

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 270-274

### **Casein kinase II phosphorylation site.**

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,  
193-197, 203-207, 287-291

### **N-myristoylation site.**

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA  
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCCT  
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG  
GCTCGTGGCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCCTCACCTGG  
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC  
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCTGCACCGGCGTCA  
GTGACTGCTCTGGGGGAAGTGAACAAGAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA  
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA  
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG  
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTACCTCTCTCAGGAATGCC  
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTGCGGAATGCCACATCCTC  
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC  
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCTGGCTCCGAGCCAGGAGCGC  
CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC  
CTCGCTGCCCTTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA  
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT  
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAAGTGGCCCTGGAGATTGAGGGTCCC  
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACTGCCACAGCCAGAACTGAG  
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG  
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

## **FIGURE 46**

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSSGSCPPTKFCR  
TSGLCVPLTWRCRDLDCSDGSDEEEECRIEPTQKGQCPCPPGLPCPCTGVSDCSGGTDKKL  
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT  
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL  
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

### **Signal sequence:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 230-246

### **N-glycosylation site.**

amino acids 126-130, 195-199, 213-217

### **Casein kinase II phosphorylation site.**

amino acids 84-88, 140-144, 161-165, 218-222

### **N-myristoylation site.**

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,  
224-230, 230-236, 263-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 44-55

### **Leucine zipper pattern.**

amino acids 17-39

## FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG  
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA  
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG  
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG  
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT  
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG  
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG  
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA  
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA  
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG  
TACCCCGACCTACACAATATGGTTTTATTATGTCGCGATGATGGAACGTGGAATAATCTGCC  
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC  
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT  
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCGGTG  
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGTCT  
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT  
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTTCTTC  
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA  
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC  
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAG  
TTCCAGCAGTGACCCTGACTTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG  
ACGAAGCTGTGAGTGGCGGCTTGAGTGCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG  
GGCTGCCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA  
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA  
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA  
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT  
GTTGTTCTTAAGAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA  
ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAA  
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC  
CTGTTTCTTCTTGACACAGACTGATTAAAAAATTAAAGNAAAAAA

## **FIGURE 48**

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI  
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI  
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS  
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP  
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT  
WPSTHETLLTTWKIVAFTATSVLLVLLLVLARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD  
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDS  
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVLF LRN

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 325-344

### **N-glycosylation site.**

amino acids 104-108, 134-138, 192-196

### **Casein kinase II phosphorylation site.**

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,  
364-368, 380-384, 467-471, 468-472

### **N-myristoylation site.**

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,  
478-484

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405



## FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA  
TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCCGAGCCCGCTCCACCGCCGT  
AGCGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG  
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC  
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA  
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT  
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA  
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA  
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT  
CACAATTTAGGAACTGGTATGTGGATGAGCCGTCTGCGGCAGCGAGGTCTGCGTGGTCTATG  
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGA  
CCGGTGCAACATGAAGAACAATTTCAATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCCTT  
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG  
GAAGAAGATGCCAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT  
CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT  
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC  
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA  
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCAATTCAGAGTGTGTT  
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA  
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA  
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG  
GTTATTAGGACATATAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC  
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT  
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT  
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA  
GAGTCTCAATAAATGTCACCTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG  
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAACACCTCTGTTTTCTTGCTCTATACAG  
CAGCACATATTATCATAACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG  
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA  
GCAGGAAAAA

## **FIGURE 50**

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLLRGGQPVCRRGGTQRPCYKVIYFHDTSRRL  
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGD F WIGLRRREEKQSNSTACQDL  
YAWTDGSISQFRN WYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY  
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLLV  
VTTVVCWWVICRKRKREQPD PSTKKQHTIWPSPHQNSPDLEVYNVIRKQSEADLAETRPDL  
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES  
GWVENEIYGY

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 117-121, 312-316

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 296-300

### **Casein kinase II phosphorylation site.**

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,  
299-303, 306-310, 323-327

### **N-myristoylation site.**

amino acids 18-24, 37-43, 76-82, 146-152

[illegible]

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC  
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT  
GTTTGCCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT  
GTTTACGATGCGCTTGTGGACCCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC  
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC  
CTGCTGAAGTTGAAAATGGTGCAGGTCTGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC  
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACCTC  
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT  
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT  
GCAGCAAATGTTTGCCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC  
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTTCGTTCCACTAACATTTTTTCGGAATCTG  
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTGAGAAAGAAGGACCCATCATCAT  
CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA  
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG  
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT  
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG  
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG  
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC  
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG  
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT  
GTTGACCTGACCATGGAACCTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA  
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT  
TCTTGAATGCCATGTGAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA  
ACTCAGGTGATGGAAGTTGGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT  
TTAAAATAAAGTGCCTTTTATACAATG

## **FIGURE 52**

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMOVVFRHGAR  
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL  
TKVGMQQMFALGERLRKKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE  
GPITIIHTDEADSEVLYPNYQSCWSLRQTRGRRRQTASLQPGISEDLKKVKDRMGIDSSDKVD  
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFHLHILES  
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMLTLGIFDHKWPPFAVDLTMELYQHLESKEW  
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSP EKYHALCSQTQVMEVGNEE

### **Signal sequence:**

amino acids 1-23

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 218-222

### **Casein kinase II phosphorylation site.**

amino acids 87-91, 104-108, 320-324

### **Tyrosine kinase phosphorylation site.**

amino acids 280-288

### **N-myristoylation site.**

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

### **Amidation site.**

amino acids 216-220

### **Leucine zipper pattern.**

amino acids 10-32

### **Histidine acid phosphatases phosphohistidine signature.**

amino acids 50-65

U  
S  
P  
A  
T  
E  
N  
T  
O  
F  
F  
I  
C  
E

Figure 1 displays 12 histograms, labeled  $k=0$  through  $k=11$ , showing the distribution of the number of non-zero elements in the vector  $x_k$ . The x-axis represents the 'Number of non-zero elements' (ranging from 0 to 10), and the y-axis represents the 'Frequency' (ranging from 0 to 10). The distributions are roughly bell-shaped and centered around 5-6 non-zero elements, with a slight shift to the right as  $k$  increases.

CTCCTCTTAACATACTTGCAGCTAAAACATAAATATTGCTTGCTTGGGGACCTCCTTCTAGCCT  
TAAATTTTCAGCTCATCACCTTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT  
GCCATTTGCACCAGACCTGGATTCCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT  
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG  
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC  
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCTCATCCA  
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG  
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCAGAGAGCTCTTTCTCCCA  
GTCCAGAGGGTGTCAAGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA  
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCGCGCAAAGTG  
TGTGCCCGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC  
TATGGCCGAAAACCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA  
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTG  
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG  
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA  
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA  
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTTCGTTGCTCAGGGGAGGAG  
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA  
TGTGGCTGTCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA  
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT  
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCTTCAGAGTTGG  
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA  
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT  
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA  
CACCATTTGTCTCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA  
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG  
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC  
TATGTGTTCAAA



[illegible][illegible]

## **FIGURE 56**

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA  
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL  
VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIISSIQKMSI  
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT  
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 104-120, 278-292

### **N-glycosylation site.**

amino acids 228-232

### **Glycosaminoglycan attachment site.**

amino acids 47-51

### **Casein kinase II phosphorylation site.**

amino acids 135-139, 139-143, 253-257

### **Tyrosine kinase phosphorylation site.**

amino acids 145-153, 146-153

### **N-myristoylation site.**

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

### **Amidation site.**

amino acids 265-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 6-17

03500401-01101  
TOTAL "671101"



## FIGURE 57

CCACGCGTCCGCTGGTGTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA  
AAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC  
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTCTAA  
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA  
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG  
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT  
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG  
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA  
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA  
GGCATTTCCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG  
CTGGACATGTCTCGGTCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA  
TTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAACAACATG  
TCTGTGTCCTAATTTCTGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA  
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG  
ATTTTTATTCCATCTTCTATAGCTTTTTTAAACAACATTGGAAAGGATCCTTCCTGAGCGTTT  
CCTGGCAGTTTTTAAACGAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA  
AAGCGCAATAAGCACCTAGTTTTTCTGAAAAGTATTACCAGGTTTAGGTTGATGTCATCTA  
ATAGTGCCAGAATTTTAATGTTTGAAGTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA  
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT  
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA  
CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA  
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT  
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA  
TGAAGGACTATATCTAGTGGTATTTACAATGAATATCATGAACTCTCAATGGGTAGGTTTC  
ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT  
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG  
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA  
AA  
AA

U.S. GOVERNMENT PRINTING OFFICE

[illegible]

**Signal sequence:**

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

[illegible]

AAAATATTATATATAAAAGTAAAAAAAAAA

## **FIGURE 60**

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL  
VGPFPGNLNPKSYAGFLTVNKTYNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH  
GPYVVTSMNLTDRDFPWTTLTSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF  
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG  
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVGTG  
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT  
EIMNMYKVLINQGLDIIVAAALTEKSLMGMDWKGSQEQYKKAEEKVWKIFKSDSEVAGYIRQ  
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

### **Signal sequence:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 81-85, 132-136, 307-311, 346-350

### **Casein kinase II phosphorylation site.**

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,  
353-357, 424-428

### **Tyrosine kinase phosphorylation site.**

amino acids 423-432

### **N-myristoylation site.**

amino acids 22-28, 110-116, 156-162, 232-238

### **Serine carboxypeptidases, serine active site.**

amino acids 200-208

### **Crystallins beta and gamma 'Greek key' motif signature.**

amino acids 375-391

## FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTGGCTACAACAT  
TTTTCCCTTTCCTAACAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT  
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG  
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG  
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC  
CTGGGGGAGGGCCTGCCTAACAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT  
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG  
TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA  
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA  
GACCAAAATAGATAACAGGATTCCTGAACATTCTAAGAGGGAGAAAGTATGTTAAAAATA  
GAAAAACCAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC  
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTT  
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT  
CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG  
ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCAGCAGCCGGC  
ATGCCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT  
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA  
ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG  
CCCCTCATCGTGACGCCCTGCAGCGAAGTGCTCACCCCTACCAACAATGTCAACAAGCTGCT  
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA  
AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCACAAGAAGGAGCACTAC  
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA  
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA  
GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT  
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT  
CTTCTACATCTACGGCTTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA  
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG  
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC  
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC  
AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG  
CAGTATCACCACCCGCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT  
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC  
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA  
CTGGACATCAACCAGCCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC  
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG  
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC  
ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTAACTATAG  
GCAACTTTATTTTCTTGGGGAACAAAGGTGAATGGGGAGGTAAGAAGGGGTAAATTTTGTG  
ACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA  
TTTCAATATTTCCCAAACCTTTAAGAAAAAATTTAAGAAGGTACATCTGCAAAAGCAA

[illegible]

**Signal sequence:**

**Transmembrane domain:**

N-glycosylation site.

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,  
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

## FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCAGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGATGACCA  
CGTGGAGCCTCCGGCGGAGGCCGGCCCGCACGCTGGGACTCCTGCTGCTGGTCTGCTTTGGGCTTCTGGTGTCTCC  
GCAGGCTGGACTGGAGCACCTTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA  
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT  
GGAGGGACCGCCTGCTGAAGATGAAGCCCTGTGGCTTGAACACCTCACCACCTATGTTCCGTGGAACCTGCATG  
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCCTTCGTCTGATGGCCCGCAGAGATCG  
GGCTGTGGGTGATTCTGCGTCCAGGCCCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCAGCTGGCTAC  
TCCAAGACCCTGGCATGAGGCTGAGGACAACCTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC  
TGATGTCCAGGGTGGTGGCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG  
GTTCTTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACCTGC  
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT  
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGAATGAGT  
ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT  
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG  
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG  
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG  
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC  
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCCT  
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG  
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT  
ACACCGTGTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG  
GCTTAATTGGAATCTCTATCTGAATGATTCACCCCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA  
GCTTCTTTAGAGGTTCCGCCCTGGACAAATGGNGTTCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGG  
GTAGCTTGTCCATCAGCTCCACGCCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATTCA  
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA  
GCAGCGGAATCAACCAGGTATCGTTTTTGAGGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC  
ACCTGGGCAGGAACCAGTACATTAAGTGAAGCGGTGGCACCCCTCCTGCTGGTGGCAGTGGGAGACTGCCGCCTC  
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAGCATCTCCTTAAGTAGCAACCTCAGGG  
ACTGGGGGCTACAGTCTGCCCCGTCTCAGCTCAAACCCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC  
TGGCTTTGTTGATGATGGCTTTCTACAGCCCTGCTCTTGTGCCGAGGCTGTCCGGCTGTCTTAGGGTGGGAGC  
AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGACCCGGACGTCACAGCCC  
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCTTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT  
TTTATCCCCGAAATCCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT  
CTTCCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACCTCGGCGTGAGAAACATGTGACTTCCCCTT  
TCCCTTCCCACTCGCTGCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC  
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCA  
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCC  
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA  
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG  
ACAGAAGGCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC  
GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTCTGGGGCCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA  
GTTGCAGTAAAGCTATAACCTTGAATCAAA

U.S. GOVERNMENT PRINTING OFFICE

## **FIGURE 64**

MTTWSLRRRPPARTLGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW  
IFGGSIHVFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDFSGNLDLEAFVLMA  
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ  
YKRGGPPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKGLSKGIVQGVLAT  
INLQSTHELQLLTTFLFNVQGTQPKMVMYWTGWFDWSGGPHNILDSSEVLKTVSAIVDAGS  
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAHYMKLRDFFGSI SGIP  
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE  
TSITSSGILSGHVHVRGQVFVNTVSI GFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN  
IDDQRKGLIGNLYLNDSPKKNFRIYSLDMKKSFFQRFGLDKWXSLEPETPTLPAFFLGSL SIS  
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSGGINQVIVFEETMAGPA  
LQFTETPHLGRNQYIK

### **Signal sequence:**

amino acids 1-27

### **Casein kinase II phosphorylation site.**

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

### **N-myristoylation site.**

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,  
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586



## FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC  
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG  
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC  
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC  
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA  
GGCAGACACTCGGTTCGTTCTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCC  
CGTTCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC  
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACCTA  
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA  
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG  
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC  
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC  
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTAGGTGGAGAATGAATATGGTAGC  
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG  
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG  
GACTCTATAACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT  
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA  
CTGGGGGCCAGAATCACTCCACACGGTCTGTGTGAGCTGTAAACCAAAGGACTAGAGAACATGC  
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG  
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC  
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTTTGCTCTTCGAGATGTCATCAGCAAGT  
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG  
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCCAT  
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC  
GAACCTATATGACCCATAACCATTTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC  
CATGACCGTGCCATATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG  
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAACTGGATATCTTGGTGGAGAACATGG  
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG  
GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG  
GTGGTTTCCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT  
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG  
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC  
ACAACAGACCCTCTACGTGCCAAGATTCTGCTGTTTCCCTAGGGGAGCCCTCAACAAAATTA  
CATTGCTGGAAC TAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC  
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACTAGG  
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC  
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTGAGGACTTCAAGA  
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG  
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC  
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA  
GACACTCCATCTCAAAAAAAAAAAAA

## **FIGURE 66**

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRGHDRFLLDGAPFRYVSGSLHYFRVPRVL  
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNNGSRDLIAFLNEAALANLLVILRPGPYI  
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE  
YGSYRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF  
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMVMFHGGTNF  
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMML  
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN  
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP  
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSGDFTFLYL  
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD  
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

### **Signal sequence:**

amino acids 1-27

### **N-glycosylation site.**

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 4-8

### **Casein kinase II phosphorylation site.**

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,  
603-607, 644-648

### **Tyrosine kinase phosphorylation site.**

amino acids 191-198

### **N-myristoylation site.**

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,  
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

## FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC  
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT  
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT  
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT  
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT  
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG  
AAAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTTCATG  
CTGTGCGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC  
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC  
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA  
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA  
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG  
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT  
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC  
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACCTCAGA  
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT  
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT  
GACTTGTTTAAAATTATGGCATAACAAAATTGTTACTATTCTCCCTCTATTACCCATGTCA  
AAAACCTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT  
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA  
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC  
CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC  
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG  
CTTGGAACCGCTGCCAGCCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGCTTGTTG  
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA  
AATATTCCCTTTGCAAATGGGATTTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC  
AATTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG  
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT  
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG  
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTTAAATTGTTTGTAACCTGGAT  
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT  
TTCTTACTAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPPLKEYSFEKVREESSFSDIPDVKNDF AFL  
LHMVDQYDQLYSKRFGVFLSEVSENKLR EISLNHEWTFEKL RQHISRNAQDKQELHLFMLS G  
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH  
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESRLRLHLKILHVKS NLTKVPSN  
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAEELQNC ELERIPHAIFSLSNLQELDLKS  
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLES LYFSNNKLES LPVAVFSLO  
KLRLCDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP  
EKVGQLSQLTQLELKGNC LDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP  
FANGI

### **Signal sequence:**

amino acids 1-20

### **N-glycosylation site.**

amino acids 241-245, 248-252, 383-387

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 326-330

### **Casein kinase II phosphorylation site.**

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

### **Tyrosine kinase phosphorylation site.**

amino acids 349-355, 375-381

### **N-myristoylation site.**

amino acids 78-84, 124-130, 212-218, 392-398

Downloaded from www.sciencedirect.com

## FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCCTTTTCATTGACAAACTGACTTTTTTTTATTTCT  
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTACACACATAAGGAT  
CTGTGTTTGGGGTTTCTTCTTCTCCCTGACATTGGCATTGCTTAGTGTTGTGTGGGGAGGGAGACCACGTGG  
GCTCAGTGCTTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC  
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTGCTGCTCTGTCTTTACTTCAAAATACACAAC  
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG  
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCTGCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC  
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA  
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG  
CTCCCATCAGTTTTCATGGAAAATAACTCAGTGCCCTGCTGGGAACCACTGCTGGAGATCCCTACAGAGAGCTTC  
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCACTCA  
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCCTGGCTGGAACCTGACGTTTCCCTGGAGGTGTCCAGAAA  
GCTGATGTAACACAGAGCCTATAAAAGCTGTGCGTCTTAAGGCTGCCAGCGCCTTGCCAAAATGGAGCTTGTA  
AGAAGGCTCATGCCATTGACCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT  
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC  
AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAGCATCAGAATTATCTTTTCTATGTCCAGCTT  
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCCTCTGCTAGGG  
CAAGTCTGCAGTAAAAACGACTATGTTCCCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT  
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCTAACATCTCTATTCCAACTGT  
GGCGGTTACCTGGATACCTTGAAGGATCCTTACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT  
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAACTAAACTTCAAAGAGATTTTCTTAGAAATAGAC  
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT  
GGCCGTGTGACTCCACCTTGAATCGTCATCAAACCTCTCTGACTGTGCTGTGTCTACAGATTATGCCAATTCT  
TACCGGGGATTTTCTGCTTCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT  
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA  
GACCCAACTTGCAGACCAAATTTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA  
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC  
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA  
GAAGATGATGTAATACAAAGTCAAAATGCCTGGGCAAAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA  
TTTGAAGAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACCTCTTTTGTTCAGTTAGTCTGCAC  
ACCTCAGATCCAAATTTGGTGGTGTCTTGTGATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC  
TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAACTTGTAAAGGTGTATCCCTTATTTGGACACTATGGGAGA  
TTCCAGTTTAAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAGTTTTGATATGTGAT  
AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG  
AAAACAGATTCCATCATAGGACCCATTCTGCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTTCAGCAT  
GAAACACATGCGGAAGAACTCCAAACAGCCTTTCAACAGTGTGCATCTGTTTTCTTTCATGGTTCAGCTCTG  
AATGTGGTGAAGTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG  
CAGAACTATTAACCTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT  
GGCTACACATATTATGAATAAATGAGGAAGGGCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

## **FIGURE 70**

MELVRRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI  
ERPENKSIRIIFSIVQLDPDGSCSENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSTLT  
FQIVTDSARIQRTVFVFFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV  
EKDYKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVL  
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIIISKSYLEAFNSNGNNLQLKDPTCRP  
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGNST  
VEIIYITEDDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN  
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL  
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDIIIGPIRLKRDRSASGNSGFQHETHA  
EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

### **Signal sequence:**

amino acids 1-24

### **Transmembrane domain:**

amino acids 571-586

### **N-glycosylation site.**

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,  
394-398, 419-423

### **Casein kinase II phosphorylation site.**

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,  
408-412, 463-467, 520-524, 556-560

### **Tyrosine kinase phosphorylation site.**

amino acids 172-180, 407-415, 407-416, 519-528

### **N-myristoylation site.**

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

U  
S  
P  
A  
T  
E  
N  
T  
O  
F  
F  
I  
C  
E

## FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG  
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGCTGCTG  
TGCTGCCGCGCCGCGCCGTGCCCTGCCACAGCGCCACGCGCTTCGACCCACCTGGGAGTCC  
CTGGACGCCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG  
GGGAGTGTTTTCCGTGCCAGCTTCGGTAGCGAGTGGTCTGGTGGTATTGGCAAAAGGAAA  
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTTCAAATATGAAGAT  
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTTCAGGC  
CTCTGGTGCCAAATACATTGTCTTAACCTTCCAACATCATGAAGGCTTTACCTTGTGGGGGT  
CAGAATATTCGTGGAACCTGGAATGCCATAGATGAGGGGCCAAGAGGGACATTGTCAAGGAA  
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTTGA  
ATGGTTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG  
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATATCAGCCTGAGGTTCTGTGG  
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT  
ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA  
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA  
CATAAATGGGAAAACCTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG  
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG  
GAGGAAATCTTTTGATGAATATGGGGCCACACTAGATGGCACCATTCTGTAGTTTTTTGAG  
GAGCGACTGAGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA  
TACCTGGCGATCCAGAAATGACACTGTCAACCCAGATGTGTGGTACACATCCAAGCCTAAAG  
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCACATCAGGACAGCTGTTCTTGGCCAT  
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAACTACTGGGCCATGGACAGCCACTTAACTG  
GATTTCTTTGGAGCAAAATGGCATTATGGGTAGAAGTGGCACAGCTAACCATTCATCAGATGC  
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG  
ATGCTGCAAGTTATGTCTAAGGCTAGGAACATATCAGGTGTCTATAATTGTAGCACATGGAGA  
AAGCAATGTAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCACTA  
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC  
TCTTCACATTGATTTGTTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTTCACATTATAGTAG  
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG  
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT  
TTTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA  
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG  
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG  
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG  
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA  
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA  
GTTTTTCTTTCTTCAATTATAAATAACATAAGTGTACTGTAACTTTACAAACGTTTTTAATT  
TTTAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAACATAAACTCATTGTGCAA  
ATGTAA

## **FIGURE 72**

MRPQELPRLAFPLLLLLLLLLLLPPPPCPAHSATRFDPWESLDARQLPAWFDQAKFGIFIHWG  
VFSVPSFGSEFWFWYWQKEKIPKYVEFMKDNYPSPSKYEDFGPLFTAKFFNANQWADIFQAS  
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLEW  
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY  
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI  
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT  
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI  
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

### **Signal sequence:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 171-175, 239-243, 377-381

### **Casein kinase II phosphorylation site.**

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,  
375-375

### **Tyrosine kinase phosphorylation site.**

amino acids 361-369, 389-397

### **N-myristoylation site.**

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

### **Leucine zipper pattern.**

amino acids 410-432

### **Alpha-L-fucosidase putative active site.**

amino acids 283-295

109604060  
TET/070707



## FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT  
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC  
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG  
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAAGAGCT  
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC  
AGCTGAATTTACAGAAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG  
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGGAGA  
TGGATTTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG  
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT  
ACTTGGAATAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA  
AACTGCAACACAAACAACAGAATTTATTGTGAGTGACAGTACCTACTCGGTGGCATCCCCCTT  
ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG  
AGAAAAAAATTGATTTGTGTACAGAAAGTTTTTATGGAACTAGCACCATGTCTACAGAAAC  
TGAACCATTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC  
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC  
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT  
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA  
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA  
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC  
TTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT  
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC  
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAGCTTCCCACGGCC  
TTTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA  
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG  
TGGGTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGAC  
CCTTCTTTCAGCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCTTCTGAGCCCGGTA  
AGAGCAAAAAGTGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG  
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT  
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA  
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT  
AGGAAATATACTTTTACAAGTAACAAAAATAAAACTCTTATAAATTTCTATTTTTATCTGA  
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA  
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT  
TGAATATTATTCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTTCA  
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTTATTTTGTGCTGAGACTAATCTT  
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG  
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA  
GCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA  
AGCATTTAGAAACTT

090401.07101

## **FIGURE 74**

MARCFSLVLLLTISIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR  
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQF  
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP  
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFF  
GAAAGLGFCYVKRYVKAFFFTNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP  
SKTTVRCLEAEV

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

### **Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

### **Tyrosine kinase phosphorylation site.**

amino acids 79-88

### **N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

Downloaded from www.sciencedirect.com

## **FIGURE 75**

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTTCGGTGCCGCGACTTTCACGATGG  
CTCGCCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA  
ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG  
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA  
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT  
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC  
TGATGACGTGCAAACCCCCCTATATATGGGCCCCTGAGTATATCAAGTACTTCAATGATAAA  
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA  
TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAC  
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC  
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA  
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG  
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA  
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGA  
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT  
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG  
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG  
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAAGTGTTCCTGAGCAAGAAAGAGATC  
TCATAGGACGGAGGGGAAATGGTTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC  
AGCTATTCAGACATCTCCATGGTTTTCTCCATGAAACTCTGTGGTTTTCATCATTCCTTCTTAG  
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA  
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC  
TTCATTAAAAGTATAAGCCTAACTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG  
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA  
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC  
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT  
TCCTTTTG

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

amino acids 1-48

## amino acids 111-125

amino acids 165-169, 185-189

amino acids 154-158, 265-269

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

amino acids 188-194, 225-231

## amino acids 244-253

## FIGURE 77

GGACAGCTCGCGGCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG  
GGGCCCCAGCCTGGCCCCGGGTCAACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA  
TTGCTCCTGCTGCCCGGCTCCTACGGA CTGCCCTTCTACAACGGCTTCTACTACTCCAACAG  
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG  
TGGTGGAGACACCCGAGGAGACCCTGTTACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC  
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAAGCT  
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT  
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG  
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA  
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA  
ACGGGCGCTACCAGTTCAACTTCCACGAGGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG  
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG  
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCGGCAGCCCTGCGGTGGCC  
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCCTGCACCGCTATGAT  
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCCTGAGAAGCT  
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC  
AGCTCTTTGCCGCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT  
GGCAGCGTCCGCTACCCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT  
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTTACGGTGTTTACTGCTACCGCCAGC  
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT  
CGTTTTCCCTTGTGGGTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT  
TTAAACATTTTTTTTACTATTTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG  
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG  
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGGCCAGAGTGGGC  
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCT  
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC  
CACCTCTACTTCTCTGTGAAGCCGCTGACCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA  
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAACTCCCCTCCCCGTTCCCCT  
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTGTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG  
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGTATGACTGCCTTCCGCCAA  
AA  
AA

0950401.07101

## **FIGURE 78**

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGHGKDLLNGVKLVVETPEETLFTYQ  
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLAIGLRHRSFGDYQGRVHLRQD  
KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ  
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR  
HRR LHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR  
CDAGWLADGSVRYPVVHHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

### **Signal sequence:**

amino acids 1-17

### **Casein kinase II phosphorylation site.**

amino acids 29-33, 53-57, 111-115, 278-282

### **Tyrosine kinase phosphorylation site.**

amino acids 137-145

### **N-myristoylation site.**

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

U  
S  
P  
A  
T  
E  
N  
T  
O  
F  
F  
I  
C  
E

## FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC  
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC  
CAGGGGAGGGTGCACCAGGCGGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA  
CTTCCAGTACGACCATGAGGCTTTCTGAGCGGGAAGTGGCCAAGGAATTCGACCAACTCA  
CCCCAGAGGAAAGCCAGGCCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC  
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA  
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG  
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCCGGTGAAGAATTTTCATGAC  
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC  
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG  
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA  
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA  
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG  
ATGGGCACCTGGATGGGAGTGAAGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC  
CTGGTGGAAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA  
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG  
ACCTGACCCGGCACCCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG  
CACAATGACCGGAGGAGGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGGAG  
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCCTGGGCTCTCAGGGACCCCCCTGGGTCTGGCTTC  
TGTCCTCTGTCACACCCCCCAACCCAGGGAGGGGCTGTATAGTCCCAGAGGATAAGCAATAC  
CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA  
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT  
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA  
AAGCCAGCGCCGGGACCTTGAAA  
AAAAAAAAAAAAAAAAA

U.S. GOVERNMENT PRINTING OFFICE

[illegible]

**Signal sequence:**

N-glycosylation site.

**Casein kinase II phosphorylation site.**

N-myristoylation site.

**Endoplasmic reticulum targeting sequence.**

amino acids 325-330



## FIGURE 81

GGGGCCTTGCCCTTCGCACTCGGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG  
GCGGCGGGGCGCGGGTGCAGGGGATCCCTGACGCTCTGTCCCTGTTTCTTTGTGCTCCAG  
CCTGTCTGTGCTCGTTTTTGGCGCCCCCGCTCCCCGCGGTGCGGGGTTCACACCGATCCTG  
GGCTTCGCTCGATTTGCCGCCGAGGCGCTCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG  
CGGGTCGTCTGTGCTCTCTCTCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCAGGGGCTCT  
GAGGAGGTGACGCGCGGGGCTCCCGCACCTTGGCCTTGCCCGCATTCTCCCTCTCTCCAG  
GTGTGAGCAGCCTATCAGTCACCATGTCGCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG  
TGTCTGCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGGGAGCCGCTCCCATTTGCTATCACATG  
TTTTACCAGAGGCTTGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCAGGGGGCTGCC  
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG  
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC  
TGGTCGAGAAAACATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT  
GGTCTGCTTCTTTACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA  
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGG  
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC  
GCCGATTTAATTTACAGAAGAATTTTGTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA  
GAAGGACCACATGTGGGCCTTGTTCAGCCAGTGAACATCCCAAATAGAAATTTACTTGAA  
AACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGTA  
ATTCGAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA  
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTATTGATGGTTGGCCTTCTGATGACAT  
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA  
AGCCTATCCCTGAAGAACTGGGGATGGTTCCAGGATGTCACATTGTTGACAAGGCTGTCTGT  
CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAA  
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT  
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTGGAGATAGCAATTTCCGCCTC  
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT  
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACA  
AAGAGAATGTCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT  
GATGCCATTTCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGGAGCCCCAACAAGAA  
CTTCTAGTAATTGTACAGATGGGCAGTCTATGATGATGTCCAAGGCCCTGCAGCTGCTG  
CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG  
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTACAGGATT  
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCAGCAAT  
AATGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT  
CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACCTATTAAGTATGTCAAC  
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT  
ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA  
GATAATGTGGATTAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA  
TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAA

## **FIGURE 82**

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEF SVY  
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASF TVT  
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDC KADIAFLIDGSFNIGQRRFNLQKN  
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFY LKNFTSAKDVLFAIKEVGFRGGNSNTGKAL  
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG  
MVQDVT FVDKAVCRNNGFFSYHMPNWF GTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI  
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI  
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAH DAGITI  
FSVGVAWAPLDDLKDMASKPKESHAF FTREFTGLEPIVSDVIRGICRDFLESQQ

### **Signal sequence:**

amino acids 1-24

### **N-glycosylation site.**

amino acids 100-104, 221-225

### **Casein kinase II phosphorylation site.**

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,  
425-429, 478-482, 528-532

### **N-myristoylation site.**

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,  
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

### **Amidation site.**

amino acids 145-149

bioRxiv preprint doi: <https://doi.org/10.1101/071001>; this version posted July 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC  
GGCGGCCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGGCCGCGAGCGCAACTCGGTCCAGTCG  
GGGCGGCGGCTGCGGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC  
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGGCCGCTCCGACGCGGACCTCGGCTCCAGTC  
AAGCCCCGGCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA  
GGTTGAGGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTAT  
CACAATGAGACCAACACAGACACGAAGTTTGAAATAATACCATCCATGTGCACCGAGAAAT  
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG  
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC  
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT  
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA  
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG  
TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA  
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG  
GAGCCTTGGAACGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG  
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC  
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG  
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT  
GCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA  
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA  
TCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC  
TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC  
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG  
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG  
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAA  
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCCTTCCATGGGCATAGGTAAGCTG  
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCC  
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC  
CCTCTCTCAGCACAGCCTGGGGAGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG  
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT  
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCACCAA  
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGAGGCATGCACATCTGGAATTAAG  
GTCAAACATAATTCTCACATCCCTCTAAAAGTAACTACTGTTAGGAACAGCAGTGTTCTCAC  
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT  
TGCAATAGTAACCTTGAAAGGTATATGACTGAGCGTAGCATAACAGGTTAACCTGCAGAAACA  
GTAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT  
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTAGGTGTCA  
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA  
TAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC  
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 84**

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ  
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG  
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMMLCTRDSECCG  
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL  
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV  
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEI

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331

### **N-myristoylation site.**

amino acids 202-208, 217-223

### **Amidation site.**

amino acids 140-144

U99004611001

## FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCTCAGAAGGACTCTG  
AAGATAACAATAATTTTCAGCCCATCCACTCTCCTTCCCTCCCAAACACACATGTGCATGTACACACACACATACA  
CACACATACACCTTCTCTCTTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTCATAGAAAAGGACAC  
TAAAGCCTTAAGGACAGGCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTGAGTCAAAAAACATGGGAGGGG  
CCAGGCACGGTGACTCACACCTGTAATCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTGAGGAG  
TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC  
AGGTGCCTGTAATCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT  
CAGCTGAGTGCACCGCTGCACTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGA  
GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACCTCTGCATCCTCTTCTTCCAGGGCTGCCCTGATGGGGCCTG  
GCAATGACTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG  
GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCCTTGGGCCCTGGCAGGGTTGCTGACCCCTACCCTGCAAAACACA  
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACCATGAGGCTTCTCGTGGCCCC  
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCTCA  
GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCGCTCGTCTACCGCGAGGCTACCCTGTGGACTGCAATGA  
CCTATTCTGACGGCAGTCCCCCGGCACTCCCCGAGGCACACAGACCCCTGCTCCTGCAGAGCAACAGCATTGT  
CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCAGAACAGCTTTTCGGA  
TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT  
GGAGGACCACAGCTTTGACAGGGCTGGCCAGCCTACAGGAACCTATCTCAACCACAACCAGCTCTACCAGCATCGC  
CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAG  
CCGCTGGTTTGAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT  
GAACTTCCGGCCCCCTGGCCAACCTGCGTAGCCTGGTGTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC  
CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACAGCTGGCCCCGGGTGCCAGGCGGGCACT  
GGAACAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGC  
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT  
GAACCTCCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCTTTCATCCACCCCCGCGCCTTCCA  
CCACCTGCCCCAGATGGAGACCCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGCAACCAGCAGACGGTGGAGTC  
CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTATCCGCTGGGCCAATGC  
CACGGGCAACCGTGTCCGCTTCATCGAGCCGCAATCCACCCTGTGTGCGGAGCCTCCGGACCTCCAGCGCCTCCC  
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGGCCCTCATCTCCCCACGAAGCTTCCCCCAAG  
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGTGCTGCAATGCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG  
GGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCAGGAGGTACCGGGTGTACCCCGAGGGGACCCT  
GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCAGAACCTGGTGGGGGCTGACAC  
TAAGACGGTTAGTGTGGTTGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGGCTGGAGCTCCG  
GGTGCAGGAGACCCACCCCTATCACATCCTGCTATCTTGGGTACCCCCACCAACACAGTGTCCACCAACCTCAC  
CTGGTCCAGTGCCTCCTCCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCGCCTGCCTCGGGGAACCCACAGCTA  
CAACATTACCCGCCTCCTTCAGGCCACGGAGTACTGGGCCTGCCTGCAAGTGGCCTTTGCTGATGCCACACCCA  
GTTGGCTTGTGTATGGGCCAGGACCAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCTTGGGCTCAT  
TGCCATCCTGGCTCTCGCTGTCTTCTCCTGGCAGCTGGGCTAGCGGCCACCTTGGCACAGGCCAACCAGGAA  
GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTCTTGGGGCTGGAGTGCCCCCTTCTGTCCGGGTGT  
GTCTGCTCCCCTCGTCTGCTGCCCTGGAATCCAGGGAGGAAGCTGCCAGATCCTCAGAAGGGGAGACACTGTTGCC  
ACCATTTGTCTCAAAATTCCTTGAAGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTTACCAA  
AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCTGGCAGCTGGGC  
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA  
CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCCATCTT  
CTCTCTGCCAGAGGCTCCTGGGCCTGGCTTGGCTGTCCCTACCTGTGTCCCCGGGCTGCACCCCTTCTCTTCT  
TCTTTCTCTGTACAGTCTCAGTTGCTTGTCTTGTGCCCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC  
CTCGGGGGGCTGCCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCAGGAA  
CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA  
ATGTGTACCTCCCCCAACCCGATTCACTCTTTTCTCCTGTTTTGTAAAAAATAAAAAATAATAATAACAATAAA  
AAAA

[illegible]

**Signal sequence:**

**Transmembrane domain:**

**N-glycosylation site.**

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

**Casein kinase II phosphorylation site.**

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,  
243-247, 313-317, 488-492, 700-704

**Tyrosine kinase phosphorylation site.**

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

## FIGURE 87

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT  
CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA  
CATCAAGTTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCCACCC  
CCTGGCCACACTCTTCAAGATCCTGGCGTCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA  
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG  
CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA  
GCGCTTCGCCGTCTTCTGTGCGGAGGTGAGTGAGAAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT  
GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT  
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCGCC  
CAGCATTGCCAGCTCACGGGCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT  
GGCCTTCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA  
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACCAACCGCTACATCGTCATCGA  
CGGGCTGCGGGAGCTCAAACGCCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCA  
AGATGTGGGCGTGCACCTGCAGAACTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAA  
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT  
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCCAGCACCT  
GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCA  
CCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGCAAGCTGCG  
CTACCTGGACCTCAGCCACAACAACCTGACCTTCTCCTGCGGACATCGGCCTCCTGCAGAACCTCCAGAACCT  
AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT  
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG  
CAACCGGCTGGAGTGCTGCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA  
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCCTGAGCGGAG  
GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCGGAGGGGAGGCCCTAGCTTCTCCAG  
AACTCCCGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGAGCTGGGGCCGCTTGTGAGTCAGGCCAGAGCAGA  
GGACAGTATCTGTGGGGCTGGCCCCCTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG  
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG  
CTGCCACCAGAGGTCTGGGACCCCTCACTTTAGTTCTTGGTATTTATTTTTCTCCATCTCCACCTCCTTCATCC  
AGATAACTTATACATTCCCAAGAAAGTTACAGCCAGATGGAAGGTGTTACAGGAAAGGTGGGCTGCCTTTTCCCC  
TTGTCTTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCAG  
CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCCTCCAGCTGGA  
AAGGCCAGGCCTGGAGCTTGCTCTTCAGTTTTTGTGGCAGTTTTAGTTTTTTTGTTTTTTTTTTTTAAATCAAA  
AAACAATTTTTTTTTTAAAAAAGCTTTGAAAAATGGATGGTTTGGGTATTAAAAAGAAAAAAGAACTTAAAAAA  
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTTCCCTTGAGCAAGCAGCCAGACGT  
TGAAGTGTGTTTTCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGGATCTGGTGTGACCTTGGTCCAGGAGTT  
CTATTTGTTTCTGGGAGGGAGGTTTTTTTTGTTTGTTTTTTGGGTTTTTTTTGGTGTCTTGTTTCTTCTCCTCC  
ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG  
ACTCGGGTTGGCTAATCCCCGGATGAACGGTGCTCCATTGCGACCTCCCCTCCTCGTGCCTGCCCTGCCTCTCCA  
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCAGACTTTGTTTCCCCACCTCCTGCGGCATGGGTGTGT  
CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTTCATGAAGAGCAGACACTTA  
GAGGCTGGTGGGAATGGGGAGGTGCCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTCCCGTCCCTGGCGC  
CTGGAGTGACACAGCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT  
AGAAGGGTCCCCGCCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT  
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTCGTGTATTGGATATAATCCTCAGAAATAATGCACACTAG  
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACCTGTAGACTCGGTACAGTATCAAATAAA  
ATCTATAACAGAAAAA

U.S. GPO: 1975 O-344-101

[illegible]

**Transmembrane domain:**

N-glycosylation site.

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

**Casein kinase II phosphorylation site.**

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447



[illegible]

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT  
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCAATTGACT  
GGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG  
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCTGGTCAT  
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAACTTTGAGGAAATTGGGC  
CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT  
GTGGATAATCCCGTGGGCACTGGGTTTCA GTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA  
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG  
AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC  
ATTGGTCTAGAGCTTTATAAGGCCATTACGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT  
TGCCTTGGGTGATTCTGATCTCCCCTGTTGATTCTGGTGCTCTCCTGGGGACCTTACCTGT  
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA  
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT  
GATCATTGAACAGAACACAGATGGGGTGAACCTTCTATAACATCTTAACTAAAAGCACTCCCA  
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC  
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA  
GCTCAAAATTATTCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA  
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC  
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTGAGGAGGCCTG  
GGTGCGGAACTGAAGTGGCCAGAACTGCCTAAATTCA GTCTAGCTGAAGTGGAAGGCCCTGT  
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC  
TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT  
GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT  
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT  
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT  
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAACCTAAGATTTTTTAAAAAATTGAT  
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

[illegible]

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,  
187-193, 195-201, 331-337, 332-338, 360-366

[illegible][illegible]

## **FIGURE 92**

MGARGALLLALLLARAGLRKPESQEAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW  
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI  
YLSPLYLGNSPYDIALVKLSAPVITYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP  
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDI FGDMVCAGNAQGGKDACFGDSSGGPLACNKNG  
LWYQIGVVSWSGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL  
LGPV

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 167-171, 200-204, 273-277

### **Casein kinase II phosphorylation site.**

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

### **N-myristoylation site.**

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,  
259-265, 269-275

### **Amidation site.**

amino acids 33-37

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 252-263,

### **Serine proteases, trypsin family, histidine active site.**

amino acids 78-84

## FIGURE 93

U950401.1 - U950401.1

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT  
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGAGCCCGACCAGCGGAGGACGC  
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT  
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTGCGGATCCCAG  
CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC  
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGCGAGCCGGAGCCAGAAAGTGCCATTCT  
GTGATCACACAGGACTTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC  
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC  
ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT  
TTTCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGAGGTGACAGGGACTGTAGGCCT  
GCATCTGGGGGTAAACCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG  
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCAGTTCTTGAGCAGTATTTCCATGAC  
TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC  
CCGTGTGGTTGGACAACAGGGGCCGGGGCCGGGCGGGATTGAGGCCAGTCTAGATGTGCAGT  
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG  
GGACAGGAGCCCTTCCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT  
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA  
AACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCCTCAGGTGACAGT  
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG  
CCCCATGTCAACACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCAAAATGAAA  
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG  
GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCAGTCCACCATCCAGTTACTTCAATGC  
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTGAGCAACA  
GAGTGCCCATTTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTGGGGGGATCCTA  
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAAG  
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC  
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCTGGCTGGGATCCTGTAACAGGC  
TGGGGAACACCAACTTCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCCTATC  
AGGAGAGATGGCTTGTCCCTGCCCCGGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG  
GAAGCCCTGCTGAACCCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA  
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT  
CCTGCCTTAGATTCTCAATAAGATGCTGTAAGTACTGATTTTGAATGCCTCTCCCTCCGC  
ATCTCATCTTTCTCTTTCAATCAGGCTTTTCAAAGGGTTGTATACAGACTCTGTGCACTA  
TTTCACTTGATATTCAATCCCCAATTCAGTGCAGGAGACCTCTACTGTACCCGTTTACTCT  
TTCCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTGTCTTTATG  
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACCTCTCTG  
ACTACTCTTGTCTTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC  
TCCATTTGTAGATTTTGTCTCTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA  
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA  
TGTAACAAA

[illegible]

**Signal sequence:**

N-glycosylation site.

**Glycosaminoglycan attachment site.**

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,  
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,  
521-527, 533-539, 549-555

[illegible][illegible]

## **FIGURE 96**

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS  
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS  
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG  
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA  
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD  
AQPASGSGSVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW  
IKGNYLDCREG

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 93-97, 207-211

### **Glycosaminoglycan attachment site.**

amino acids 109-113, 316-320

### **Casein kinase II phosphorylation site.**

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

### **N-myristoylation site.**

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

### **Serine proteases, trypsin family, histidine active site.**

amino acids 171-177

05904011-01101  
TOTTO "01101"



## **FIGURE 97**

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGGTGGTTT  
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG  
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA  
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA  
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC  
ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACTGTTCTCTGTGCTGCTGGG  
GGCCTGGCAGCTGGGGAACCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC  
CCCACCCTGTGTATTCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG  
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT  
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC  
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT  
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT  
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG  
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC  
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA  
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG  
CCGCGCGCTCCTTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA  
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCGCGGTAAATAGGCTCATCTACC  
TCTACCTCTGGGGGCCCCGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCCGAC  
GGCCTCAGGCCCCCCTCCAAGGCATCAGGCCCCGCCAACGGCCTCATGTCCCCGCCCCAC  
GACTTCCGGCCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT  
AGGTATTTGTAACCCTGCCACATATCTTATTTATTCTCCAATTTCAATAAATTATTTATT  
CTCCAAAAAAAAA

090001.071

## **FIGURE 98**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP  
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA  
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPNTHCWISGWGSIQDG  
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC  
QVDGAWLLAGIISWGECAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG  
SGAAARS

### **Signal sequence:**

amino acids 1-32

### **N-glycosylation site.**

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### **Glycosaminoglycan attachment site.**

amino acids 826-830

### **Casein kinase II phosphorylation site.**

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

### **Tyrosine kinase phosphorylation site.**

amino acids 607-615

### **N-myristoylation site.**

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

[illegible]

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA  
CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGTATGGT  
GGAGCTGCACAACCTCTACCGGGCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA  
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC  
AACAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCCGCATCACAGACGAGGGCATGGACGT  
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT  
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC  
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT  
GGTGTGCAACTATGAGCCTCCGGGGAACTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC  
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC  
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA  
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT  
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCC  
CCAACCTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC  
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA  
CCTTCCCCAAATCGACCCATGTTCCCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA  
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG  
GGAACCTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG  
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC  
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA  
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCTTGCCAGGTGCAGAGGGCCCTGACA  
AGCCTAGCGTTGTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA  
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTTGAATGGGATACCACTCAAAGGG  
TGAAGAGGTCAGCTGTCCTCCTGTCATCTTCCCCACCCTGTCCCAGCCCCATAACAAGATA  
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC  
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC  
GGGCCCACACCTCTCCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT  
CACTGCCTACCTGGCCTGGGGCTGTCTGCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG  
TGTAAGCTGGGGATGGGGATTCTTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC  
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT  
GTCCAACCTGTGAAA

## **FIGURE 100**

MHGSCSFLMLLLPLLLLLLVATTGPVGALTDEEKRLMVELHNLVYRAQVSPTASDMLHMRWDEE  
LAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ  
MCGHYTQVVWAKTERIGCGSHFCEKLQGVETNIELLVVCNYEPPGNVKGKRPYQEGTPCSQC  
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFVLVTEV  
SGSLATKALPAVETQAPTSLATKDPSPMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS  
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAEELPPSSEVLAS  
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV  
SGLNSGPGHVWGPLLGLLLLLPPLVLGIF

**Signal sequence:**

amino acids 1-22

**N-glycosylation site.**

amino acids 114-118, 403-407, 409-413

**Glycosaminoglycan attachment site.**

amino acids 439-443

**Casein kinase II phosphorylation site.**

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

**N-myristoylation site.**

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,  
250-256

**Amidation site.**

amino acids 82-86, 172-176

**Peroxidases proximal heme-ligand signature.**

amino acids 287-298

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.**

amino acids 127-138

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.**

amino acids 160-172

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIGURE 101

GTAAGTGAAGTCAGGCTTTTCATTGGGAAGCCCCCTCAACAGAATTGCGTCATTCTCCAAGTTATGGTGGACGT  
ACTTCTGTTGTTCTCCCTCTGCTTGTCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTTCAT  
CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACTGAACAACAATGAATTGGAGACCATTCC  
AAATCTGGGACCAGTCTCGGCAAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA  
ACATCTGAAAGAGTTTCAGTCCCTTGAACTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAAACTGCATT  
TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA  
TTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACT  
GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG  
TGCTCTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAACTTATGGATGGAGCTTTTGGGGGCTGAGCAA  
CATGGAAATTTTGCAGCTGGACCATAACCAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGTGATGCT  
GCAGGAACCTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG  
TGAGCTGGACCTAACTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATAC  
ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT  
GGATCTGAAGAACAATGAAATTTCTGGACTATTGAAGACATGAATGGTGTCTTCTCTGGGCTTGACAAACTGAG  
GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA  
TCTAGACCTGAGTGACAACGCAATCATGCTTTTACAAGGCAATGCATTTTACAAATGAAGAACTGCAACAATT  
GCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAAACAACCTT  
TCAGAGCTTTGTAAATGCCAGTTGTGCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGTCTGTAGCCGAGA  
TGGCTTTGTGTGATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTT  
CAATTTGAGTTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGTCTGGAAAAAAGACAATGA  
ACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC  
CATCCTTCCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCATCTCCAATCACTTTGGTTC  
ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCAT  
CCGAGCTGGGGCCATGGCAGCTTGGAGTGTGCTGTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGA  
TGGGGGCACAGACTTCCAGCTGCACGGGAGAGACGCATGATGTGATGCCCGAGGATGACGTGTTCTTTATCGT  
GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAGCAATGC  
AACTCTGACTGTCTTAGAAACACCATCATTTTGTGGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC  
CGTCTACAGTGCAATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTTGGTGGT  
AACCAGAGAGGCATTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA  
ATACACATGTGAGATGCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC  
CTGCGACTCCCCCTCAGATGACAGCCCCCTGCTTAGACGATGACGGATGGGGCCACTGTGGTGTCTGATCATAGC  
CGTGGTTTGTGTGTGGTGGGCACGTCACTCGTGTGGGTGGTTCATCATATACCACACAAGGCGGAGGAATGAAGA  
TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCTAGTTATTTGTCTCATCTCAGGGAACGTT  
AGCTGACAGGCAGGATGGGTACGTGTCTCAGAAAGTGGAAGCCACCACAGTTTGTACATCTTTCAGGTGCTGG  
ATTTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC  
CACAGATCTGTTCCCTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA  
TCCTTTTGAACATATCATAACAGTTGTCAGTCTTGACCCAAGAACAGTTTTTAATGGACCACTATGAGCCAGTTA  
CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC  
GTGGCCTTTCATGTGAGGAAGCTACTTAACACTAGTTACTCTACAATGAAGGACCTGGAATGAAAAATCTGTG  
TCTAAACAAGTCCTCTTTAGATTTTAGTGCAATCCAGAGCCAGCGTCGGTTGCCCTCGAGTAATCTTTTCATGGG  
TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCTAGCC  
AAGAGCCTTTTATTTGAAAGCTCATTCTTCCCAGACTTGGACTTGGGTGAGAGGAAGATGGGAAAGAAAGGAC  
AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACCTACAGGACTCCAAATTTTCAGTC  
TTATGACTTGGACACATAGACTGAATGAGACCAAAGGAAAAGCTTAAACATACTACCTCAAGTGAACCTTTTATTTA  
AAAGAGAGAGAATCTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAAATGCTTTATTTATACAGAT  
GAACCAAAATTACAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTTAAACTA  
TTTTTTAACTTTGTTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT  
TTATAATGCCAGATTTCTTTTTATGGAAGTGAAGTACTAAAGCATTTTTAAATAATACCTGCCTTGTACCATTTT  
TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

U.S. Patent & Trademark Office

## **FIGURE 102**

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVVSAN  
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP  
GYFDNLANTLLVLKLNRRISAIPPKMFKLPLQHLLELNRNKIKNVDGLTFQGLGALKSLKM  
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE  
FCQKLSELDLTFNHL SRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNE  
ISWTIEDMNGAFSGLDKLRRLLIQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ  
MKKLQQLHLNTSSLLCDCQKWLQPQWVAENNFSQSFVNASCAHPQLLKGRSIFAVSPDGFVCD  
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENY AHLRAQG  
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA  
MARLECAAVGHPAPQIAWQKDG GTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN  
SAGSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER  
HFFAAGNQLLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG  
WATVGVVIIAVVCCVVGTSVLVWVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTLD  
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP  
MYLKGNVYGSDPFETYHTGCSPDPRTVLMHDHYEPSYIKKKECYPCSHPSEESCERSFSNISW  
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDA  
YSSFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEEHNICTFKQTLNRYRTPNFQS  
YDLDT

### **Signal sequence:**

amino acids 1-19

### **Transmembrane domain:**

amino acids 746-765

### **N-glycosylation site.**

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### **Glycosaminoglycan attachment site.**

amino acids 826-830

### **Casein kinase II phosphorylation site.**

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

### **Tyrosine kinase phosphorylation site.**

amino acids 607-615

### **N-myristoylation site.**

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

## FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG  
AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC  
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA  
TAAACCAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAAATAAATAATCTTAAAGCAGCTGTTCCCTCC  
CCACCCCCAAAAAAGGATGATTGGAATGAAGAACCGAGGATTACAAAGAAAAAAGTATGTTCATTTTCTC  
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT  
GGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAA  
TTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA  
TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTCTTCATCAACCTCCTTTTTTTTAAAT  
TTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACCACCTGGATTTCATCTGGATGTTGCT  
GTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATGTTGAACAAGAT  
GCGAATACATCCACAGCAGATAATGATAGGTCTAGGTTTAACAGGGCCCTATTTGACCCCTGCTTGTGGTGCT  
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCAGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  
CCAGTTTCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT  
GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT  
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTTGGGGCTTTCAATGGTCTGGCGAAACCTCAACACTCTGGA  
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT  
GCGAAACAACCCCATTTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCCTTCTTTCGCGCCGACTAGACTTAGG  
GGAATTGAAAAGACTTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTTGAGGTATTTGAACCTTGC  
CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAGCTGGATCTTTCTGGGAATCA  
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAAGTGTGGATGATACAGTCCAGAT  
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC  
ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG  
TAACCTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTGTGCGCCGGTG  
TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT  
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCCTCCACATC  
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCAGGATAGCTGT  
GCTCAGTGATGGTACGTTAAATTTACAAATGTAACCTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA  
TTCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC  
AACCCTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCCTCC  
AGTGGTTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTTCGACAGAGAAAACCTT  
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTGATGAAGACTACCAAATCATCAT  
TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA  
TCGGCAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC  
CATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT  
CAACCACACAACAACAGTTAAACACAATAAATTCATACACAGTTTCAGTGCATGAACCGTTATTGATCCGAATGAA  
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAACATTTACAGAGTTACAAAAACAAACAATCAAAAAA  
GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAA  
AAAAGAAAAGAAATTTATTTATTAATAAATTTCTATTGTGATCTAAAGCAGACAAAA

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG

## **FIGURE 104**

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC  
VRKNLREVPDGISTNTRLNLHENQIQIIKVNSFKHLRHLLEILQLSRNHIRTIEIGAFNGLA  
NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS  
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLELDLSGNHLSAIRPGSFQGLMHLQKL  
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL  
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE  
LKCRASLTLSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTQDTGMYTCMVSNVSGN  
TTASATLNVTAATTTTFFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSITPQ  
STRSTEKFTTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQON  
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS  
VHEPLLIRMNSKDNVQETQI

### **Signal sequence:**

amino acids 1-44

### **Transmembrane domain:**

amino acids 523-543

### **N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,  
442-446, 488-492, 606-610

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

### **Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

### **N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,  
391-397, 422-428, 433-439, 531-537



# FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTTCCCTCCTGCTGTTTGGGGGCA  
 TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC  
 GAGGGCGGGCGTGCACCTTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCGCGGGCTGGGAGCTTCGGGTAGA  
 GACCTAGGCCGCTGGACCCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCGCGGGGTGGGGCTGCTGCTGTGC  
 GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGCGGTGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC  
 GAGCGCCCATGCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGCTCT  
 CCCGAGCCACTCCCGTCTGGGTGCTCGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC  
 ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA  
 GTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG  
 TTTCACTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTGCATTTCCAGCCCTACAG  
 CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA  
 CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCAACTGCCAA  
 CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT  
 CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG  
 CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTTCAT  
 CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA  
 ACTTTCATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG  
 AACACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC  
 AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC  
 CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT  
 GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAACTGCAACAATTGCATTTAAATACA  
 TCAAGCCTTTTGTGCGATTGCCAGCTAAAAAGGCTCCACAGTGGGTGGCGGAAAAACAACCTTTCAGAGCTTTGTA  
 AATGCCAGTTGTGCCCATCCTCAGCTGTAAAAAGGAAGCATTTTGCTGTAGCCAGATGGCTTTGTGTGT  
 GATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC  
 ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGCTTGGAAAAAGACAATGAACTACTGCATGAT  
 GCTGAAATGGAAATATGACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG  
 CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCAATCACTTTGGTTTCATCTACTCTGTC  
 AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCATGGATCTCACCATCCGAGCTGGGGCC  
 ATGGCACGCTTGGAGTGTGCTGTGTGGGGCACCCAGCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC  
 TTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGACA  
 GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAGCAAATGCAACTCTGACTGTC  
 CTAGAAACACCATCATTTTTCGGGCCACTGTTGGACCGAAGTGAACCAAGGGAGAAACAGCCGCTCTACAGTGC  
 ATTTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC  
 TTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG  
 ATGTCTAACACCCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGATCCCCACTCCAACCTGCGACTCCCCCT  
 CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGTGT  
 GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC  
 AACACAGATGAGACCAACTTGGCAGCAGATATTCCCTAGTTATTTGTCTCATCTCAGGGAACGTTAGCTGACAGGCAG  
 GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACAGTTTGTACATCTTCAGGTGCTGGATTTTTCTTACCA  
 CAACATGACAGCTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTT  
 CTTTGTCCGTTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGAAATGTGTATGGCTCAGATCCTTTTGAACA  
 TATCATACAGGTTGCAGTCTTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCAGTTACATAAAGAAAAAG  
 GAGTGCTACCCATGTTCTCATCCTTCAGAAAGATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT  
 GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC  
 TCTTTAGATTTTAGTGAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA  
 GCTCTCAGGAGACCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGATCCCCACTCCAACCTGCGACTCCCCCT  
 TTGAAAGCTCATTCTTCCCGAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGACAGATTTTCAGGAA  
 GAAATCAGATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC  
 ACATAGACTGAATGAGACCAAAGGAAAAAGCTTAACTACTACCTCAAGTGAACCTTTTATTTAAAGAGAGAGAAT  
 CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAAATGCTTTATTTATACAGATGAACCAAATTAAC  
 AAAAGTTATGAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTTAACTTTG  
 TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGATTTTTTATAATGCCAGA  
 TTTCTTTTTATGGAAATGAGTTACTAAAGCATTTAAATAATACCTGCCTGTACCATTTTTTAAATAGAAGTT  
 ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

U990401.07

## FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR  
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT  
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY  
FDNLANTLLVLKLNRRNRI SAIPPKMFKL PQHQHLELNRNKIKNVDGLTFQGLGALKSLKMQR  
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC  
QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS  
WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK  
KLQQLHLNTSSLLCDCQLKWL PQWVAENNFSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF  
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENY AHLRAQGGE  
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA  
RLECAAVGH PAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA  
GSISANATLTVLETPSFLRPLLDRTVT KGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF  
FAAGNQLLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDD GWA  
TVGVVIIIAVCCVVGTS LVVWVIIYHTRRRNEDCSITNTDETNPADIPSYLSSQGT LADRQ  
DGYVSSSESGSHHQFVTSSGAGFFLPQHDS SGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY  
LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCHPSEESCERSFSNISWPS  
HVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDAYS  
SFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEEHICTFKQTL ENYRTPNFQSYDLDT

**Signal sequence:**

amino acids 1-27

**Transmembrane domain:**

amino acids 808-828

**N-glycosylation site.**

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,  
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

**Glycosaminoglycan attachment site.**

amino acids 886-890

**Casein kinase II phosphorylation site.**

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,  
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,  
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,  
1073-1077, 1079-1083, 1081-1085

**Tyrosine kinase phosphorylation site.**

amino acids 667-675

**N-myristoylation site.**

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,  
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

**Leucine zipper pattern.**

amino acids 58-80, 65-87

U55444 "04444"

# FIGURE 107

CAAACCTTGCGTCGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC  
 TGGGGGAGCGCGTTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC  
 CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCTGGGTGGTCCCGTCCCTATCCCTCCTTTATATA  
 GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG  
 GCGCACAGCATTCCGAGTTTACAGATTTTACAGATACCAATGGAAGGCGAGGAGGCAGAACAGCCTGCCTGGT  
 TCCATCAGCCCTGGCGCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCCAGAGCCGGGTGCTGC  
 TGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA  
 GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG  
 AGCCCGGGCCTGGCCCAGCCCGGGTCAGCTGCCCCGAGACTGTGCCTGTTCCAGGAGGGCGTCTGTGGACTGTG  
 GCGGTATTGACCTGCGTGAGTTCCCGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACCAACCAGC  
 TGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACACTGAACCTGCAAAAACAACCGCCTGA  
 CTTCCCGAGGGCTCCAGAGAAAGCGTTTGGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC  
 TGACCTTGGCACCCCGCTTCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT  
 ATGGGCTCACCTTTGGCCAGAAGCCAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC  
 TGCCGGACAACATGTTCAACGGCTCCAGCAACGTGAGGTCTCATCTGTCCAGCAACTTCTTGCGCCACGTGC  
 CCAAGCACCTGCCGCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT  
 TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA  
 CTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCAGCTGGGCTGC  
 CGCGCAGCCTGGTGTCTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCATCC  
 GCAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCCACTGGCCTTCCAGGGCC  
 TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTGGCCAGTGGCCTGCCTCGCCGCGTGC  
 GCACCCTCATGATCTTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTTGGAGG  
 AGCTCAACCTCAGCTACAACCGCATCACAGCCACAGGTGCACCGCGACGCTTCCGCAAGCTGCGCCTGCTGC  
 GCTCGTGGACCTGTGCGGCAACCGGCTGCACAGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGTGCAAGG  
 TCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA  
 CCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA  
 TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAA  
 AGATTAGTGCGGTGCCCCGCAATGCCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC  
 TGGCTGTGGGCTCCGTGGTGGACAGTGCCCTTCGGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT  
 TAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAAGAGGAGGAGGAGG  
 AGGAAGAGGAAACAAGATAGTGACAAGGTGATGCGAGATGTGACCTAGGATGATGGACCGCCGACTCTTTTCTGC  
 AGCACACGCCCTGTGTGCTGTGAGCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA  
 TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCACGGCGTGTCCACGGCCAGACACATGC  
 ACACACATCACACCCTCAAACACCCAGCTCAGCCACACACAACCTACCCTCAAACCAACCACAGTCTCTGTACAC  
 CCCCCTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCCCTGCCCTGGCACACACAGGCACCCA  
 TTCCCTCCCCCTGTGACATGTGTATGCGTATGCATACACACCACACACACACATGCACAAGTCATGTGCGAA  
 CAGCCCTCCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT  
 GTCCATCTGTCCGTCCGTTCCCTGGAGAAGACACAAGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCCTCT  
 GGAACCTCAAAAAGCTGGCTTTTATTCTTTCCATCCTATGGGGACAGGAGCCTTACAGGACTGCTGGCCTGGCC  
 TGGCCACCCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA  
 CAGGCACTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG  
 CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTTT  
 GTTCTTACAGGCCTGTGGGGGAAGTTCCGGGTGCCCTTTATTTTATTCTTTTCTAAGGAAAAAATGATAAAAT  
 CTCAAAGCTGATTTTTCTTGTTATAGAAAACTAATATAAAAGCATTATCCCTATCCCTGCAAAAAA

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808

**Signal sequence:**

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,  
535-557

## FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG  
CAGTAGCCCCGGGCGGCGAGGGCTGGGGGTTCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACCACC  
CAACCTGTTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGT  
GCTGGTATCTCTCGCTCTACCTGTCAGGCGGCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGAT  
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC  
TGTGTGCCAACCCAGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG  
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA  
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG  
CTCCATGGCAAACCTGTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT  
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT  
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTATAAAGGCTTCGATCTCATGTATATTGGAGG  
CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA  
CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA  
AGTTATGATTGAACCTTCAGGTCCAATTTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA  
TAATAATTGGATTCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA  
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCTACTCCACCACCACCACCACC  
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACACCACCGGACTGACAACCTATAGC  
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCCTCAGAAACCCAGAGG  
AGATGTGTTTCTGTTTCTGGTACACAGTGTGAATTTTGAACATGGACTTTGTGGATGGATCAGGGAGAAAGACAA  
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTGCGGCAGCCAAAGCCCCAGG  
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTAGGGGACCTGTGCCTGTCATTAGGCA  
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAAACCGGTGCCCACGGAGCAGCCCTGTG  
GGGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA  
AAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAATAAAGGAACTGGGATTATTGAGCCTGGAGAAG  
AGAAGACTGAGGGGCAAACCATTTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG  
TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG  
GGCCATTGTTAGAATACTTCATAAAAAAAGAGTGTGAAATCTCAGTATCTCTCTCTTTCTAAAAAATTAGA  
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAA  
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG  
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAATGCTGAGCTCACTGAAATA  
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAAATGAGAGTATGATGGAC  
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC  
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTCAGAGAG  
ATTTTCATCGGGTGCACTCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC  
ACACCGGCAGACCTTTCTTTCACCTCATCAGTATGATTAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC  
AGAACAGTAATATTTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTAACCTGGTAAAGGCAGGGCTGG  
AGGGGGAAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTTATT  
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAAATAAATGAAAATTTTACTTTTCGATGCCAA  
TGATACATTGCACTAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT  
CTAAAAATAAAAAATGTTAGTGGTTTTTCCAAATGGCCTAATAAAAACAATTATTTGTAAATAAAAAACTGTTAGTAAT

U.S. GOVERNMENT PRINTING OFFICE: 1967

## **FIGURE 110**

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP  
RCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD  
GSCSSALTCSMANCQYGC DVVKGQIRCQCPSGLHLAPDGRTCDVDECATGRASCPRFRQC  
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQG  
DGLTCVYIPKVMIEPSGPIHVPKGN GTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP  
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN  
RVQTDPOKPRGDFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG  
KAARLVLP LGRMLHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI  
TLRGADIKSESQR

### **Signal sequence:**

amino acids 1-17

### **N-glycosylation site.**

amino acids 273-277

### **Casein kinase II phosphorylation site.**

amino acids 166-170, 345-349

### **Tyrosine kinase phosphorylation site.**

amino acids 199-206

### **N-myristoylation site.**

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,  
421-427, 433-439, 462-468, 476-482

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 104-116, 186-198, 231-243

### **Cell attachment sequence.**

amino acids 382-385

### **EGF-like domain cysteine pattern signature.**

amino acids 75-87

## FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTGCCCCCTTTAGATTGTGA  
AATGTGGCTCAAGGTCTTCACAACCTTTCTTTCTTTGCAACAGGTGCTTGCTCGGGGCTGA  
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC  
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA  
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCCTGACTTGGAATACC  
AACACAAGTTTACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT  
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACCTATCTGCCAGTCA  
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTTCATCCTCCCT  
CTGGGGCTGTGGAGTATGTGGGGAACATGACCTTGACATGCCATGTGGAAGGGGGCACTCGG  
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC  
TCCCCAAAACAATAACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT  
GCCTGGTGAGGAACCTGTGAGTGAATGGAAAGTGATATCATTATGCCCATCATATATTAT  
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT  
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT  
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT  
GCATCTGAGAAAGTAGCCGAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC  
CGGCAGGCAAGATGAACTCATTTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG  
CACAGAAAGGAAAATCATTGTACCTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATT  
ATATCCATGTGTCTTCTCTTCTATGGAAAAATATCAACCCTACAAAGTTATAAAACAGAA  
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG  
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG  
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT  
GTATGAAGTTATTCAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACCTTTCATGG  
GCTAAACAGTACATTCGAGTGAAATTCCTGAAGAAACATTTTAAGGAAAAACAGTGGAAGT  
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA  
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT  
GTGCAACAGAAAAACATGTTGGGGAAATATTCTCAGTGGAGAGTCGTTCTCATGCTGACGG  
GGAGAACGAAAGTGACAGGGGTTTCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA  
ATTAGTTCTACTCTACACTTTCACATATCATCAACACTGAGACTATCCTGTCTCACCTACAAA  
TGTGGAACCTTTACATTGTTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG  
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA  
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACCTATGCCTTCTCTTTTTTTTCAATCACC  
AGTAGTATTTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTTA  
TTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAAAAAAAAAAA

**THE UNIVERSITY OF CHICAGO**

**Signal sequence:**

**Transmembrane domain:**

N-glycosylation site.

**Casein kinase II phosphorylation site.**

**Tyrosine kinase phosphorylation site.**

N-myristoylation site.

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18



GCAAGCGCGCAAATGCGGCCCTCCGGGAGTCTTGACAGTTCCCCTGGCAGTCTTGTTGCTGTT  
GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA  
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCGTGGTGCCCTGCT  
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA  
TATTGCGAAAGTAGATGTACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC  
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG  
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTTATCATG  
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA  
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT  
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC  
AGATTGCCTTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATAACCATAACCCTTCAAAAAAAT  
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGAGAGGAGGAACAAGAGGCGGATGAAGAA  
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCACAGAATGC  
CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG  
TTATCTTAATATTATGATTTTGATAAAAACAGAAGATTGATCATTTTGTTTGGTTTGAAGTG  
AACTGTGACTTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA  
CATTGAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTAAAGCACAGTATGATG  
GTTTAAATAGTTCTCTAATTTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT  
TTAATAATAACCTATTTCAAGTCTGAGTTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT  
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTGTATATAATTTTTCTCTG  
TTTCACTGTGTGAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG  
AAATGTGTATTTCAAGTGACAATTTCTGTGTTCTTTTTAGAGGTATATTCCAAAATTTCTTGT  
ATTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTTCTACACA  
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAAGTTCATGGTATTCTCTTGATTC  
CAACAAAGTTTGATTTTTCTCTTGATTTTTCTTACTTACTATGGGTACATTTTTTTATTTTT  
CAAATTGGATGATAATTTCTTGAAACATTTTTTTATGTTTTAGTAAACAGTATTTTTTTTGT  
GTTTCAAACCTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAATTT  
TTGGCCACTTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT  
TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAGCCTTGGA  
TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG  
CATCTTCTTGATATATGCTTAAATGTATTTTTGTCTCATATACAGAAAGTTCTTAATTGAT  
TTTACAGTCTGTAATGCTTGATGTTTTAAATAATAACATTTTTTATATTTTTTAAAGACAA  
ACTTCATATTATCCTGTGTTCTTTCTGACTGGTAATATTGTGTGGGATTTACAGGTAAAA  
GTCAGTAGGATGGAACATTTTAGTGATTTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT  
CACCTTAAAGAAGGGGGAAATCATAAATACAATGAATCAACTGACCATACGTAGTAGAC  
AATTTCTGTAATGTCCCTTTCTTTCTAGGCTCTGTGTGCTGTAATCCATTAGATTACAG  
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT  
AAAGAGTTTTGGATGTGTAACCTTGATGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT  
TTCTAACCACTTCATTAAAGCTGAAAAAAAAAAAAAAAAAAAA

## **FIGURE 114**

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL  
QPEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI  
NFISDKEWKSIIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL  
ATLFSGLLLGLCMIFVADCLCPSKRRRPQYPYPYPSKLLSESAQPLKKVEEQEAEDEEDVSE  
EEAESKEGTNKDFPQNAIQRSLGPSLATDKS

### **Signal sequence:**

amino acids 1-26

### **Transmembrane domain:**

amino acids 182-201

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

### **Tyrosine kinase phosphorylation site.**

amino acids 107-115

### **N-myristoylation site.**

amino acids 20-26, 192-198

### **Amidation site.**

amino acids 25-29

## FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTAACTAATTCAACAAACGGGACCCTT  
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCAGTGGGACAGGCGGATTGGAAGAGCGGG  
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC  
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG  
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC  
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC  
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAAACTGG  
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCAGGACTCAGCTGCA  
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG  
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA  
GAGGGGAACCTTCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG  
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATAACGGTGTGTGGATGGAGCAGGTGCT  
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA  
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC  
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGGAATCTGCGGTACTTTGAGCAGTTATTGGA  
GGAAGAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAGGCA  
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGGG  
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA  
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA  
TCGTACAGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA  
CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTGCCGCTA  
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC  
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTGCAAAT  
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG  
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG  
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG  
TTCTGGTACAACCTCTTGCAGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC  
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT  
TGAGACCTTGTGGATCAACAGAAGTTGACTGGACATCCTTTTCTGTCTTCCCCTTCCTGGTC  
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTTCTATCAGGCT  
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT  
GTGACTGAAGTCCAGCCCTTCCATTACAGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA  
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT  
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTTTTTAT  
GGCATTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

## **FIGURE 116**

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA  
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP  
TDEDEIGAALKMRLQDQTYRLDPGTISRGLPGTKYQAMLSVDDCFGMGRSAYNEGDDYYHTV  
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLR  
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTTPRRQKRLF  
CRYHHGNRAPQLLIAPFKEEDEWDSPIVRYDDVMSDEEIERIKEIAKPKLARATVRDPKTG  
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS  
RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGECDYR  
TRHAACPVLVGCKWVS NKWFHERGQEFLRPGSGTEVD

**Signal sequence:**

amino acids 1-17

**N-glycosylation site.**

amino acids 115-119, 264-268

**Glycosaminoglycan attachment site.**

amino acids 490-494

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 477-481

**Casein kinase II phosphorylation site.**

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,  
346-350, 365-369, 385-389, 457-461, 530-534

**Tyrosine kinase phosphorylation site.**

amino acids 71-80, 489-496

**N-myristoylation site.**

amino acids 14-20, 131-137, 171-177, 446-452

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 8-19

**Leucine zipper pattern.**

amino acids 213-235

09604031-071001

[illegible]

CAGCATATTGAGTTTTACTTCTCTCTCTTTTTAGTGGAAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT  
 TTCATTTATTACCGTTTTGGCTGGGGGTAGTTCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA  
 AGACAGGACAATCTTCTTGGGGATGCTGGTCTTGGAAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC  
 CAGGTTCTCTGGTTAAAACCTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC  
 CCTGGGGCACCCACCTGGCAGGGCTACCCACATGCGCATGAGCTCCCTGTGGCTCTGCTGCGGCCAGCGCTTC  
 CCTCATCTTAGGGCTGTCTCTGGGGTGACGCTGAGCCTCTCGGGTTTCTCGGATTCAGGCGGGAGGGAAG  
 ATCCCTGTGTGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTTCAGAGCTCGGCTAGACCAAAGTC  
 ATGAAGACTTCAAACCCCGATTGTCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC  
 GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCTGACCTCCCGAGCTACACTGTCCA  
 CTTTGGCCGTGGCTGTGAACCGTACGGTGGGCCATCACTTCCTCGGTTACTCTACTTCACTGGGCAGCGGGGG  
 CCCGGCTCCAGCAGGATGTCAGGTGGTGTCTCATGGGGATGAGCGGCCCGCTGGCTCATGTGAGAACCCCTGC  
 GCACTTACACACACTTTGGGGCCACTACGACTGGTCTTTCATCATGTCAGGATGACACATATGTGACAGGCC  
 CCCGCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAGACCTGTACTTAGGCCGGGCAGAGGAGTTCACTT  
 GCGCAGGCGAGCAGGCCCGGTACTGTCTATGGGGCTTTGGCTACCTGTTGTACGGAGTCTCCTGCTTCGTCTGC  
 GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCTGACGAGTGGCTTGGACGCTGCCTCATTG  
 ACTCTCTGGGCGTCGGCTGTGTCTACAGCACCAGGGGCAGCAGTATCGCTCATTGAACTGGCCAAAAATAGGG  
 ACCCTGAGAAGGAAGGGAGCTCGGCTTTCTGAGTGCCTTCGCCGTGCACCTGTCTCCGAAGGTACCCCTCATGT  
 ACCGGCTCCACAACCGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAATAAGAACCACTGACGGCTCAGA  
 TCCGGAACCTGACCGTGCTGACCCCGAAGGGGAGCGAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCTCTTCA  
 CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG  
 CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA  
 ATCGGCGCTATCAGCCCCGCTGCGCTTCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC  
 GGGGCATGGAGTACACCTTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGCGGGCCCTGGCTCGCA  
 GGGTCAGCCTGCTGCGGCCACTGAGCGGGGTGGAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGTGC  
 AGCTGGTGTCTGCCACTCTGGTGGCTGAAGCTGTGACGCCCCGGCTTTCTCGAGGCGTTTGACGCCAATGTCC  
 TGGAGCCACGAGAACATGCATTGCTCACCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG  
 ACCATTCTTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG  
 CTGTGCGAGCAGAGGCCCCCTTCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCTGTGGACACTCTCT  
 TCTTCTTACCACCGCTGTGGACAAGGCTGGGCCCCGAAGTCTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT  
 GGCAGGCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCTCGCCTGTCAACACAGAGATCACCCACGGGCCCC  
 CGGGGCTGGCCCTGACCCCCCTCCCTCCTGGTGTGACCCCTCCCGGGGGCTCCTATAGGGGGGAGGATTG  
 ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAAC  
 TGGCAGGCCAGGAAGAGGAGGAAGCCCTGGAGGGGCTGGAGGTGATGGATGTTTTCTCCGGTTCTCAGGGCTCC  
 ACCTCTTTCCGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCACGGCTCAGTGAAG  
 AACTCTACCACCGCTGCCGCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCAGCTGGCTATGGCTCTCTTTG  
 AGCAGGAGCAGGCCAATAGCACTTAGCCCCCTGGGGGCCCTAACCTCATTACCTTTCTTTGTCTGCCTCAGCC  
 CAGGAAGGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTAAATATGAAATGTTATTAA  
 ACATGTCTTCTGCC

## **FIGURE 118**

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPD SRARLD  
QSDedFKPRIVPYRDPNKPYPKKVLRTRYIQTELGSRRERLLVAVLTSRATLSTLAVAVNRTV  
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWWFFIMQDDTY  
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG  
DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE  
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV  
LGWDYFTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLNNGYR  
RFDPARMEYTLDLLLEC VTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLP LL  
VAEAAAAPAFLEAFAANVLEPREHALITLLL VYGPREGGRGAPDPFLGVKAAAELERRY PG  
TRLAWLAVRAEAPSQVRLMDVVS KKHVPDTLFFLT TVWTRPGPEVLNRCRMNAISGWQAF FP  
VHFQEFNPALSPQRSPPGPPGAGDPSPSGADPSRGAPIGGRFDRQASAEGCFYNADYLAA  
RARLAGELAGQEEEEALEGLEVMDFLRFSGHLHFRAVEPGLVQKFSLRDCSPRLSEELYHR  
CRLSNLEGLGGRAQLAMALFEQE QANST

### **Signal sequence:**

amino acids 1-15

### **Transmembrane domain:**

amino acids 489-507

### **N-glycosylation site.**

amino acids 121-125, 342-346

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

### **Casein kinase II phosphorylation site.**

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

### **Tyrosine kinase phosphorylation site.**

amino acids 736-743

### **N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-678

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

### **Cell attachment sequence.**

amino acids 247-250

Figure 1 consists of 12 histograms arranged horizontally, labeled  $x_0$  through  $x_{11}$ . Each histogram shows the frequency of non-zero elements in the vector  $x_k$ . The x-axis for each histogram is labeled  $x_k$  and ranges from 0 to 10. The y-axis is labeled 'Frequency' and ranges from 0 to 10. The distributions are roughly bell-shaped and centered around 5, with the peak frequency increasing from 5 at  $k=0$  to 10 at  $k=11$ .

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCTGTCCCCAAGCC  
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG  
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA  
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA  
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTA AAACC  
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG  
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG  
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACA ACTG  
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTTGTTAA  
AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT  
GTGGGTATGGAAGGAGGAATTGTCCTTAAGTG TAGAATCAATGAAAAGACTTAACAGCCTTCT  
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC  
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA  
AAAGATGTATTTAATAACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC  
CAACCAGGTAGTAGAAGGCTGTTGTT CAGATATGGCTGTTACTTTTAATGGACTGACTCCAA  
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATT TGGGCATATTTTCAAT  
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAAGTGGTAGAAAAGCG  
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA  
TACAGCTGTATGTTTTCTTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG  
TCAGTAGTACATTTTTTAAATGAGGGTGGTTTTTTTTCTTTAAACACATGAACATTGTAAATG  
TGTTGGAAGAAGTGTTTTAAGAATAATAATTTTGCAAATAAACTATTAATAAATATTATAT  
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT  
AAAAAATTTTAAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT  
TGTGATTAAAGTAAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT  
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAATGTACA ACTAAGCAACTAAAGAAA  
ATTAAAGTGAAAGTTGAAAAAT

**10**      **11**      **12**      **13**      **14**      **15**      **16**      **17**      **18**      **19**      **20**

amino acids 1-33

amino acids 121-125, 342-346

## amino acids 319-323, 464-468

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

## amino acids 736-743

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-672.

## amino acids 14-25

amino acids 247-250



## FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG  
AAAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAATGAATTCATCTAAATCAT  
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT  
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCAT  
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTACAGAGCTCTCCTGCT  
ACAATTATGGATCAGGTTCAAGTCAAGTGTGTTGTCCTTGAAGTGGGAATATTTTCAATCC  
AGCTGCTACTTCTTTTCTACTGACACCATTTCTGCGGCTTAAGTTTAAAGAACTGCTCAGC  
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCTTACAAGA  
AACCTAAAATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA  
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA  
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA  
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG  
AACAAAGGAAAATCTCTTTAAAGAACAGAAGGCACAACCTCAAATGTGTAAAGAAGGAAGAGCA  
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC  
ATAAGTATTTGTTACTCTGATACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAAA  
AAA  
AAAAA

U594403.1 - 07.1.01

## **FIGURE 122**

MNSSKSSETQCTERGCFSQMFLLTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN  
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ  
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCAITMRDSS  
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

### **Signal sequence:**

amino acids 1-42

### **N-glycosylation site.**

amino acids 2-6, 62-66, 107-111

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

### **N-myristoylation site.**

amino acids 15-21, 74-80, 155-161

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 27-38

05504011-03101  
TOTEST-03101

## FIGURE 123

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCCTCGACATGGCGCTGAGGCGGCCACCGCGAC  
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC  
TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAAGTGTCTTGCATCATTACGGATTTCGC  
AGACAAGTGACCCCAGGATCGAGTGGAAAGAAAATTCAGATGAACAAACCATATGTGTTTTTTTGACAACAAAA  
TTCAGGGAGACTTGGCGGGTCTGTCAGAAATCTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG  
ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA  
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC  
ACTGCCAGGAGAGTGAGGGCCACCCCGGCTCACTACAGCTGGTATCGCAATGATGTACCAC TGCCACGGATT  
CCAGAGCCAATCCCAGATTTTCGCAATCTTCTTTCCACTTAAACTCTGAAACAGGCAC TTTGGTGTTCCTGCTG  
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG  
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGGTTCTGGTTGTCTTGTCTACTGGCCCTGA  
TCACGTTTGGGCTGTGTCATACAGACGTGGCTACTTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  
ACCCAGGGAACAGATGCTGAGTTAACTACATCCGCATGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG  
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA  
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA  
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAGAATTTTCTCAAGATGGACCCGGTAAATATAACCACAA  
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTTCTGGCCTGATTCCCGCATGAGTATTAGG  
GTGATCTTAAAGAGTTTGCTCAGTAAACGCCCGTGTCTGGGCCCTGTGAAGCCAGCATGTTACCAC TGGTCTGT  
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGCGGGGAACCCA  
GAAAAGGCTTCTTACACAGCAGCCTTACTTTCATCGGCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC  
TGATCGGTGTGTCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAATCAGGAAG  
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCCTGAGGAACCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA  
ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTATTATTTTATAAAATTT  
TACATCTAAATTTTGTCTAAGGATGTATTTGATTATTGAAAAGAAAATTTCTATTAAACTGTAAATATATTGT  
CATACAATGTTAAATAACCTATTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT  
TGGAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCTTTCT  
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT  
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT  
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG  
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT  
GCCATCCTTGGGCCCCTGGCAGTGGCTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC  
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCAGCTTTGGGCTCCTGTAACAGACCTCT  
TTTTGGTTATGGATGGCTCAGAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTGTTTAATTATTTGTT  
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAAATGGAT  
CCCACTGTTTCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGATTTTCAAAACAAACCATGAT  
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTGAGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG  
TGAAACGCCTGAATCAAAGCAGTTTTTCTAATTTTGAATTTTAAATTTTTTATCCGCGGAGACACTGCTCCCAT  
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT  
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCCTGGA  
GAATGGCTCTCACTACTACCTTGTCTTTTCAAGCTTCCAGTGTCTTGGGTTTTTTTATACTTTGACAGCTTTTTTTT  
AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTTGCCGAGGCCCGCTGGCAGAGGCA  
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC  
CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACGCTCCAGCCTCTC  
TCTTGGTTGTCTAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAACCTTCTACACTAGTGCCA  
TGGAACAGGCTCTGAAAAAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA  
CGGAAAAGGAATACTCGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT  
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAAC  
CTCATTTATAAAAGCTTCAAAAAAACCA

TTTCTTAAAGGCTCTGC

100

[illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible]